

the CATO JOURNAL

VOLUME 38 NUMBER 1 • PUBLISHED BY THE CATO INSTITUTE • WINTER 2018

Editor
JAMES A. DORN

Managing Editor
TOM CLOUGHERTY

Book Review Editor
TREVOR BURRUS

EDITORIAL BOARD

TERRY L. ANDERSON
Property and Environment
Research Center

CHARLES W. BAIRD
California State University,
East Bay

RANDY E. BARNETT
Georgetown University School
of Law

JOHN H. COCHRANE
University of Chicago

KEVIN DOWD
Durham University

RICHARD A. EPSTEIN
University of Chicago

JAGADEESH GOKHALE
University of Pennsylvania

JAMES GWARTNEY
Florida State University

STEVE H. HANKE
Johns Hopkins University

RANDALL S. KROZNER
University of Chicago

DEEPAK LAL
University of California,
Los Angeles

STEPHEN MACEDO
Princeton University

ANTONIO MARTINO
University of Rome, Luiss

JEFFREY A. MIRON
Harvard University

KEVIN M. MURPHY
University of Chicago

GERALD P. O'DRISCOLL JR.
Cato Institute

SAM PELTZMAN
University of Chicago

JUDY SHELTON
Fredericksburg, Virginia

VERNON L. SMITH
Chapman University

JOHN B. TAYLOR
Stanford University

ROLAND VAUBEL
Mannheim University

RICHARD E. WAGNER
George Mason University

LAWRENCE H. WHITE
George Mason University

LELAND B. YEAGER
Auburn University

THE CATO JOURNAL (ISSN 0273-3072) is published in the winter, spring/summer, and fall by the Cato Institute, 1000 Massachusetts Ave., N.W., Washington, D.C. 20001-5403. Annual subscriptions are \$22 for individuals and \$50 for institutions. Single copies are \$8. Two-year and three-year subscriptions are \$38 and \$55, respectively, for individuals, and \$85 and \$125, respectively, for institutions. Foreign subscribers should add \$5 per year for regular delivery and \$10 per year for airmail delivery.

Correspondence regarding subscriptions, changes of address, acquiring back issues, advertising and marketing matters, and so forth should be addressed to the Publications Department. All other correspondence, including requests to quote or reproduce material, should be addressed to the editor.

Unsolicited manuscripts cannot be returned and will be acknowledged only if accompanied by a stamped, self-addressed envelope. There is no submission fee. Articles should focus on specific public policy issues, be scholarly, but also be intelligible to the interested lay reader. Accepted manuscripts must conform to the **CATO JOURNAL**'s style requirements. A style sheet is available on request.

Authors may submit their papers electronically, preferably in MS Word, to the editor (jdorn@catoinstitute.org), along with an abstract of no more than 250 words.

Unsolicited book reviews cannot be returned and will be acknowledged only if they are accepted for publication. They should conform to the above requirements and should provide an in-depth analysis of major studies of public policy. Reviews should range from 800 to 1,000 words.

The views expressed by the authors of the articles are their own and are not attributable to the editor, the editorial board, or the Cato Institute.

THE CATO JOURNAL is listed in *Cabell's Directory*, *Journal of Economic Literature*, *Current Contents/Social and Behavioral Sciences*, *America: History and Life*, *Historical Abstracts*, *International Political Science Abstracts*, and the ABI/INFORM database.

Printed in the United States of America.

Copyright © 2018 by the Cato Institute.
All rights reserved.

ARTICLES

Leszek Balcerowicz and Marek Radzikowski
The Case for a Targeted Criticism of the Welfare State 1

Roger Douglas and Robert MacCulloch
Welfare: Savings Not Taxation 17

Allan M. Malz
Liquidity Risk after the Crisis 35

Steven Gjerstad
Reorganization of Failing Financial Firms:
A Capital Structure Solution 65

James A. Dorn
Monetary Policy in an Uncertain World: The Case for Rules 81

Charles W. Calomiris
Reforming the Rules That Govern the Fed 109

Athanasios Orphanides
Improving Monetary Policy by Adopting a Simple Rule 139

Kevin Dowd
Against Helicopter Money 147

Mickey D. Levy
Sorting Out Monetary and Fiscal Policies 171

Colin Fristoe
Health, Sustainability, and the Political Economy of
Food Labeling 185

David T. Stevenson
A Review of the Regional Greenhouse Gas Initiative 203

Peter McAllister
A Teacher's Perspective on What's Wrong with Our Schools 225

Corey A. DeAngelis and Heidi Holmes Erickson
What Leads to Successful School Choice Programs?
A Review of the Theories and Evidence 247

Ryan H. Murphy and Robert A. Lawson
Extending the Economic Freedom of the World Index
to the Cold War Era 265

James D. Gwartney and Hugo M. Montesinos
Former Centrally Planned Economies 25 Years after the
Fall of Communism 285

BOOK REVIEWS

**China's Great Migration: How the Poor
Built a Prosperous Nation**
Bradley M. Gardner
Reviewed by Daniel Griswold 311

**Taxing the Rich: A History of Fiscal Fairness
in the United States and Europe**

Kenneth Scheve and Daniel Stasavage

Reviewed by **Ari Blask**

314

Churchill and Orwell: The Fight for Freedom

Thomas E. Ricks

Reviewed by **Matthew Feeney**

320

**The Limits of the Market: The Pendulum between
Government and Market**

Paul De Grauwe

Reviewed by **Thomas A. Hemphill**

324

**Oppose Any Foe: The Rise of America's
Special Operations Forces**

Mark Moyar

Reviewed by **Brad Stapleton**

331

Against the Grain: A Deep History of the Earliest States

James C. Scott

Reviewed by **Jason Kuznicki**

333

THE CASE FOR A TARGETED CRITICISM OF THE WELFARE STATE

Leszek Balcerowicz and Marek Radzikowski

Criticism of the welfare state would be more effective if it were better targeted. That is why we begin by clarifying the meaning of the welfare state. We then distinguish three dimensions of the welfare state and use them to show the extent of differences between welfare states. Next, we identify countries that appear to face the greatest challenges regarding the sustainability of their welfare states. We conclude with some remarks on how to make criticism of the overgrown and badly structured welfare state more effective.

Some Clarifications

It is best to define the welfare state not through its noble goals but through its instruments.

Instruments of the Welfare State

At the minimum they include social transfers in cash (e.g., PAYGO pensions, unemployment benefits, family allowances, and programs for the poor) and social transfers in kind (especially publicly financed health and education but also some programs for

Cato Journal, Vol. 38, No. 1 (Winter 2018). Copyright © Cato Institute. All rights reserved.

Leszek Balcerowicz is Professor of Economics and Head of Comparative Studies Department at Warsaw School of Economics as well as Founder and Council Chairman of the Civil Development Forum Foundation (FOR). Marek Radzikowski is Assistant Professor of Economics at Warsaw School of Economics. The authors thank Aleksander Łaszek and Rafał Trzeciakowski for their research assistance.

the poor). To these components of social spending, one can add, as two additional instruments, tax subsidies (e.g., subsidies to the employer-financed health insurance in the United States, see Feldstein 2005) and “social” regulations like the minimum wage, rent controls, and employment protection. In this article, we focus on social spending because it is by far the most important instrument of the welfare state and there is abundant data. However, occasionally we will also refer to the two other policies.

Welfare Arrangements

It is useful to regard the welfare state as a special kind of welfare system, which we define as arrangements to deal with various risks facing individuals—such as acute poverty, sickness, and accidents. A brief look at history reveals the existence of various welfare arrangements—for example, family (kin) based, religion based, civil based, corporate based, and market based (insurance through jobs, private savings, and commercial insurance). Countries have always had some type of welfare system combining all or some of the above arrangements. The Poor Laws in Europe, which constituted the early welfare states (Tocqueville [1835] 1997), revealed the incentive problems of government support and foreshadowed those of the modern welfare state.

The concept of a welfare system (as distinct from the welfare state) is a useful communication device:

- First, it highlights a basic fact that a lack of large welfare state does not need to mean the absence of the alternative welfare arrangements.
- Second, it draws the attention to the question of what happens to these arrangements when the welfare state expands and shrinks. There is a large body of empirical literature on the crowding-out effects of the growing welfare state (e.g., Morduch 1999, Kelley 1998) and, more interestingly, for the present situation of the unsustainable welfare state, there is some literature on the crowding-in effects of the shrinking welfare state (Heutel 2009).
- Third, distinguishing various welfare arrangements should lead to a more comparative analysis of the welfare state. There are many ideologically idealized descriptions of this system and not enough comparative research that would show how different

the treatment of the welfare beneficiaries is by welfare officials and other suppliers of welfare services (e.g., family, churches, other voluntary organizations, and the employers of the commercial organizations).

Criticisms of the Welfare State

There are many criticisms of the welfare state. They can be grouped into economic, moral, philosophical, and political. The distinctions between these categories are to some extent arbitrary but still useful because different types of criticism lead to different proposed remedies. Any proposal for reform should start with a clarification of what its main goal is.

The economic criticism of the welfare state ultimately focuses on its negative impact on the long-run economic growth. The intermediate variables are the tax burden, reduced private savings, reduced employment, and chronic fiscal fragility or outright fiscal crises.

The moral criticism shows how the repeated deviant behavior (using or misusing various social benefits) erodes social norms such as honesty, a strong work ethic, and family values (Lindbeck, Nyberg, and Weibull 1999; Niskanen 1996).¹ Moral criticism also highlights how some social benefits undermine the family and lead to increased criminality, especially among young men brought up without fathers.

Philosophical and political criticism refers to the erosion of social norms that are especially important for conservative critics of the welfare state. The “liberal” (libertarian) critics focus on two related problems. First, the loss of individual freedom due to the increased taxes needed to fund the welfare state; people are “forced to contribute toward the costs of some activity which does not further his interests or may even be diametrically opposed to them” (Wicksell 1967: 89). Second, the rise of interest groups supporting the welfare state weakens civil society (Murray 1984).

As we mentioned, these criticisms differ in their focus. But they have a large common ground, which makes it easier to form coalitions for reforming the welfare state. For example, reducing tax

¹This impact is largely independent of the national culture, as strong incentives, including the perverse ones, overwhelm the cultural specificities. For example, the Germans speak of *Sozialbetrug* (social cheating).

burdens can be justified on both economic and liberal (libertarian) grounds. The elimination of the perverse incentives makes both economic and moral sense.

Dimensions of the Welfare State

In mobilizing people for reforming the welfare state, we have to go beyond a general notion of the welfare state so that people can see which of its features produce the worst excesses. An elementary step in this direction is to distinguish three dimensions of the welfare state: its size, design, and financing.

Size

The size of the welfare state is usually measured by the ratio of social spending to GDP.² Size is a very important dimension because it is strongly linked to a country's tax burden and—to some extent—fiscal fragility and the frequency of fiscal crises, all of which matter for long-term economic growth (Balcerowicz and Rzońca 2015).

Design

Welfare states differ not only in their size but also in their design or structure. To be sure, structure is related to size: countries with large PAYGO pension systems and publicly funded health services tend to have large welfare states. However, they can be designed or structured in various ways that have important implications for their future growth (Feldstein 2005, Börsch-Supan 2012).

There also are important differences in the design of the less fiscally important components of the welfare state, which matter for the strength of the perverse incentives and the resulting social traps they produce. These microstructures increase the relative utility of non-working income and include the ease of access to, the duration of, and the replacement ratios of various social benefits.

The structural issues of the welfare state include the distinction between the welfare state, which consists of many social programs,

²In comparative research, one can also use the ratio of public spending to GDP as it is almost perfectly correlated to that of social spending to GDP. This correlation reveals that the cross-country differences both in public spending and in its dynamics have been caused by social spending.

and the *universal* welfare state that would include the guaranteed income proposal (a negative income tax) originally proposed by Milton Friedman (1962) and developed by Charles Murray (2006). The main argument for the latter option is that it would abolish the extended social bureaucracy and the pressures related to the possibility of lobbying for special social benefits. The conservative critics would object that social benefits should be “deserved” (i.e., at least means tested), and not evenly distributed. There are also some serious practical problems, especially regarding the transition to the system of guaranteed income in countries with large PAYGO pension systems. Finally, even if such a transformation succeeded there would still remain the issue of how to prevent politicians from reintroducing the special programs and politicizing the level of guaranteed income.

The above remarks are not meant to disqualify the guaranteed income proposal. Every reform has to meet two challenges: first, it has to be launched, and second, once launched, it must be sustained. There is no reform that would solve those two problems just because of its design; and there is no good substitute for a well-organized civic effort to reform the welfare state.

Financing

The third clarification refers to the structure of taxes necessary to finance the welfare state. All taxes harm economic growth but direct taxes are worse in this respect than indirect ones (Acosta-Ormaechea and Yoo 2012).

In assessing the impact of the welfare state on economic growth we should, therefore, consider the differences in the tax structure. In looking at this issue from the point of view of political economy, one would perhaps prefer to have a worse tax structure on the assumption that increasing direct taxes generates more taxpayer resistance than relying on indirect taxes. However, we are not able to assess this difference empirically. Besides, even if it were politically more difficult to increase direct taxes than indirect ones, the pressure for increased spending could overwhelm taxpayers’ resistance. In other words, the strategy to “starve the beast” may not be effective (Niskanen 2006). It appears again that there is no good substitute for civic efforts aimed at the source of the problem—namely, the pressure to increase social spending.

The Diversity of Welfare States

Let us now use our three dimensions of the welfare state—size, design, and the mode of financing—to show how diverse welfare states are in the world. True, many, if not most of them, suffer from serious problems, but the mix of welfare states and their intensity differ across countries. Therefore, the combinations of required solutions differ, too. Let us start with the size of the welfare state.

Differences in Size

As one can see from Figure 1, there is a huge variation in the ratio of social spending to GDP at all levels of GDP per capita. There is a weak positive correlation between these two variables, which gets stronger when we eliminate the outliers. But it is the outliers that are very interesting. The size of the welfare state among the richer countries (those with per capita GDP of Intl\$40,000 or more, based on purchasing power parity in current international dollars) ranges from 7 to 16 percent of GDP in Singapore, Taiwan, Korea, and Hong Kong, and from 40 to 45 percent in Denmark and France. Sweden spends less than these outliers thanks to reforms introduced in the 1990s.

It may come as a surprise that the United States spends more than Canada, Australia, Switzerland, and Russia and only slightly less than Israel and the Czech Republic. All these countries constitute a group of moderate spenders, but only in relative terms. Therefore, one cannot strongly object to Arthur Brooks's (2012) statement that "America is already Europe," especially when one considers the extent of occupational licensing and the blocking power of the teachers' unions in the United States.

The range of social spending to GDP among poorer countries (up to Intl\$20,000) and middle-income economies (Intl\$20,000 to 40,000) is almost as wide as among richer countries. The highest spenders in the first group are Ukraine, Serbia, and Brazil, and the lowest are Indonesia, Peru, China, and Mexico. In the second group, Portugal and Slovenia are the highest spenders and Chile and Kazakhstan, the lowest (Figure 1).

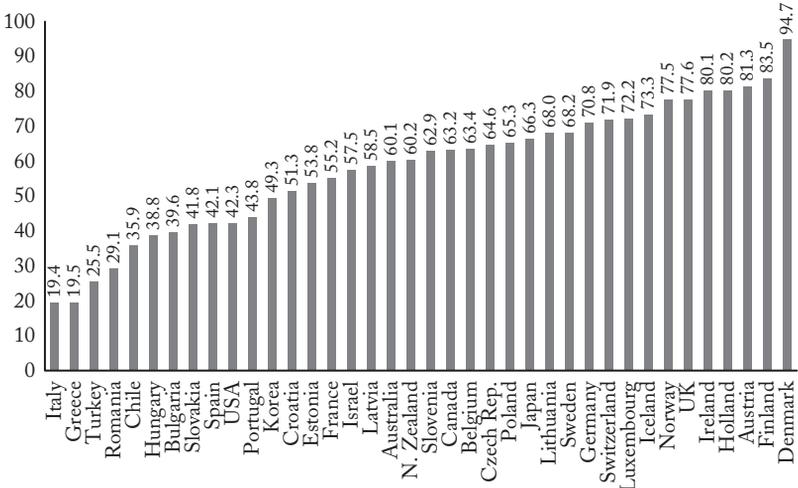
The post-socialist countries in Europe inherited a large welfare state due to a PAYGO pension system and socialized health care, and most of them have a social spending to GDP ratio of 25 percent or more. This contrasts with China, which accelerated its growth since the late 1970s, while letting its social spending decline from more than 30 percent of

GDP to less than 15 percent. It was largely following the earlier Asian Tigers’ strategy of creating welfare through economic growth and the related job creation rather than through the expansion of social spending. The opposite strategy, as practiced in some Latin American countries (e.g., Brazil and Uruguay), does not have a good track record.

Differences in Design

Design or structure is a very important dimension of the welfare state, because large structural differences bring about large differences in outcomes. However, it is impossible to rank countries according to the quality of the design of their welfare states because there is no single synthetic index to adequately measure it. Nevertheless, there is a massive literature on the perverse incentives created by the various combinations of social welfare benefits, taxes, and “social” regulations on which one can draw. For example, countries differ in the “participation tax,” which shows the fraction of any additional earnings that is “taxed away” by the combined effects of higher taxes and reduced benefits when one takes a job (Figure 2).

FIGURE 2
PARTICIPATION TAX RATES AS PERCENTAGE OF GROSS EARNINGS, 2014



NOTE: Participation in work at 100 percent of average wage from inactivity, one earner couple with two children.

SOURCE: *OECD.Stat.*

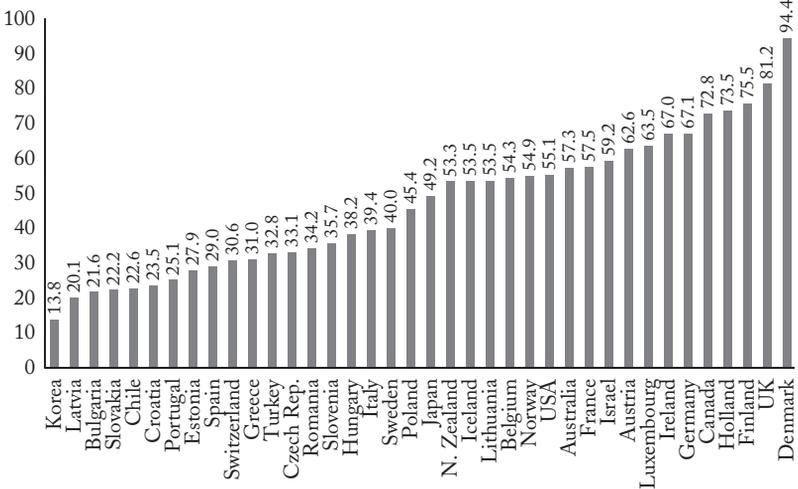
There are also huge differences in the marginal effective tax rate, which shows the fraction of any additional earnings that is “taxed away” by the combined effect of higher taxes and reduced entitlements when an individual increases the number of working hours (Figure 3).

Finally, the implicit tax on continued work for the pensioners ranges from less than zero in Poland, Iceland, and Denmark to more than 70 percent in Luxembourg and Slovenia, and more than 90 percent in Greece (OECD.Stat data).

The leftist anti-capitalists tend to deplore what they call “social exclusion,” and to blame it on free markets. In reality, blame should go to the perverse features of the welfare state, which condemn some people to inactivity and thus undermine their family life. High marginal taxes on taking on a job or on working more or longer are just one example of this category (Tanner and Hughes 2015).

Another “social trap” is created by high reservation wages resulting from excessive minimum wages or excessive levels of social

FIGURE 3
MARGINAL EFFECTIVE TAX RATES FROM
PART-TIME TO FULL-TIME WORK AS A
PERCENTAGE OF GROSS EARNINGS, 2014



NOTE: Part-time to full-time work from 50–100 percent of average wage; one earner couple with two children.

SOURCE: *OECD.Stat*.

benefits. Those instruments of the welfare state are especially destructive when combined with poor public education that prevents certain groups from permanently moving out of poverty. Moreover, if these groups happen to be minorities, extreme tension is likely to appear.

Finally, there are big differences in the tax structures among welfare states. Some countries rely more on direct taxes than on indirect ones, especially Denmark, Norway, Finland, Belgium, Switzerland, the United States, and Canada. Others, like Sweden, Greece, and the post-socialist countries, collect more indirect than direct taxes. Most of the big spenders (except Sweden) and moderate ones among the richer countries have a worse tax structure than the post-socialist economies.

A particularly bad situation exists in those countries that have big PAYGO pension systems financed by high payroll taxes (e.g., Belgium, Austria, Italy, and Greece). They combine a big welfare state with taxes that are particularly detrimental to legal employment.

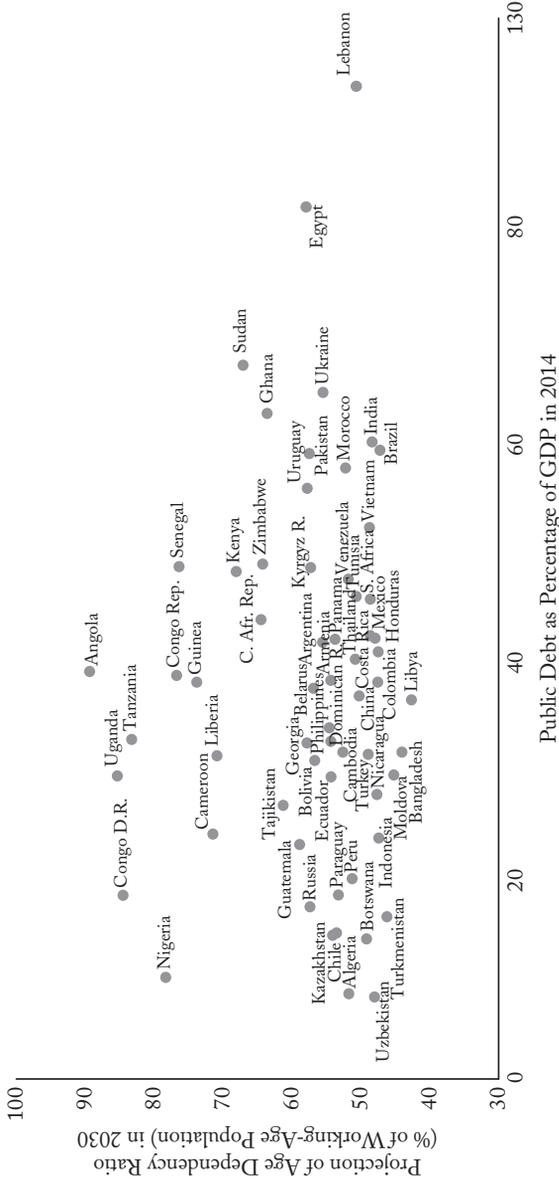
Different Challenges Faced by Different Welfare States

A small minority of countries have managed to keep a relatively limited welfare state (see Figure 1). The challenge for them is to prevent their limited welfare states from expanding. The growing problems of large (and often badly structured) welfare states should perhaps make it easier to be successful in that task. One also would hope that there would be less propaganda focusing on the “European social model.”

However, most countries in the developed world and many poorer economies face the problem of how to contain an already large welfare state and, if possible, to reverse its growth. The situation in this group differs depending on their debt to GDP ratio and on the pace of aging of their populations (Figures 4 and 5). One also would like to know the outlook for long-run economic growth, but it is very difficult to ascertain because it depends on policy choices and thus on politics, which are difficult to predict.

As one can see, the worse combinations of public debt and aging characterize Japan and Greece, followed by Italy and Portugal. On the opposite side we have Hong Kong (no public debt), Israel (good demography, moderate public debt) followed by Australia, Denmark, Switzerland, Korea (moderate aging, moderate public debt). It is striking that the Scandinavian countries appear to be in a

FIGURE 5
PUBLIC DEBT AND THE PROJECTION OF AGE DEPENDENCY
RATIO IN LESS DEVELOPED COUNTRIES



SOURCES: IMF *World Economic Outlook* and World Bank *World Development Indicators*.

better situation than Italy, Spain, France, and Germany. In addition, other large countries (the United States and Britain) are facing fiscal pressures due to a relatively high public debt.

There is a large variation among the poorer countries, too. We should remember that a lower public debt to GDP ratio among them constitutes a similar debt burden as faced by richer countries with much higher ratios of public debt to GDP. We can then see that Egypt, Brazil, Pakistan, and Uruguay face a difficult fiscal situation.

Most poorer countries (but not China) will not be burdened in the near future by a growing old age dependency ratio. In contrast, especially in Africa, they will have a growing share of working-age younger population. And, in the absence of job-creating reforms, this would lead to growing political tensions.

The solutions to the various socioeconomic problems, including those of the overgrown welfare states, are at two levels: (1) the economic level, where one must identify the best mix of policies, and (2) at the political-economy level, where one must ensure that these policies are introduced and sustained. There is no shortage of well-developed and professionally credible proposals at the first level. For example, curbing the excessive pension spending would require reforming the inherited large PAYGO systems—moving to a defined contribution system and increasing the retirement age—and raising the share of the funded system (Feldstein 2005, Börsch-Supan 2012). Similarly, there are well-thought-out solutions for the health care sector (e.g., introducing HSAs) and for unemployment benefits (e.g., the privatization proposal offered by Feldstein 2005). There is also a large body of empirical literature showing how to deal with various “social traps” (see, e.g., OECD 2014 and 2016).

Clearly, and not surprisingly, the key issue is the political economy of reforming the welfare state. We now turn to that topic in our concluding remarks.

Conclusion: The Political Economy of Welfare Reform

Policies, including institutional reforms, are actions of politicians that result from the interplay of various factors. For example, wind-fall gains in various forms (oil bonanzas and sudden reductions in the interest rates) reduce politicians’ and the public’s incentives for fiscal consolidation and encourage the growth in public spending. In contrast, some crises may even force the nonreformers in power to

do what they blocked before. Differences in the personalities of ruling politicians also matter—for example, there would have been far fewer or no fundamental reforms in Britain if it was not for Prime Minister Thatcher. There are complex interactions between situational variables, personality factors, and interest groups (Balcerowicz 2015). In the following, we will focus on the last factor.

Interest groups can be divided into statist and anti-statist. The former are driven by ideological or pecuniary motivations and aim at keeping an expanded state or even increasing the scope of interventionism. The latter, on the other hand, aim at reducing the scope of the state and are motivated by their beliefs in the value of individual freedom, the rule of law, and limited government. One of the reasons statist groups often prevail may be the simple fact that they include groups that are motivated by the prospects of pecuniary benefits (e.g., budgetary subsidies, tax preferences, and anti-competitive regulations) from expanding the size and scope of government.

However, statism does not need to prevail: much depends on the activity of the anti-statist groups and individuals. Let us, therefore, finish with some remarks on how these forces can better oppose the welfare state and press for reform. Here are some suggestions:

- Show that shrinking and restricting the welfare state leads to a better welfare system (a greater role of voluntary organizations and markets). Instead of just fighting the welfare state, libertarians and other anti-statists should fight for a better welfare system in the broad sense as discussed in this article.
- Unmask the logical deficiencies of key concepts used by the welfare statist groups such as “social rights” and “social justice” (de Jasay 2007).
- Focus on the main deficiencies of actual welfare states and not on “the welfare state” as such. Show how an overgrown and unreformed PAYGO pension system penalizes the younger generations. Show how bad public education and various social traps create the “socially excluded” groups. Show that a huge gap exists between the rhetoric of the welfare states and the results of their policies.³ Take away the “high moral ground” that the welfare statist groups claim to occupy.

³See, e.g., Goulard (2014) on the French welfare state: “The government’s priority is to defend those who are already best protected.”

- Expose the mainstream fallacy that the present welfare state results from various “market failures.” No market failure can explain the growth of the welfare state. Moreover, there are private market alternatives to a growing welfare state that are crowded out by that growth.
- Focus on the doctrines and resulting policies that obstruct reforms of the welfare state. One obstruction is the inequality debate, which confuses inequality of opportunity with that of income, and confuses fighting poverty with fighting inequality. A further impediment to reform is the revival of crude Keynesianism in the guise of an anti-austerity doctrine. Finally, unconventional monetary policy, by fueling asset bubbles, has contributed to the inequality of wealth, thus strengthening the political pressures for increased redistribution.
- Present and unmask policies that favor the rich such as restrictive zoning, crony capitalism, and insufficient competition. Demand the elimination of these privileges in the name of libertarian egalitarianism.

These are just a few examples of how to make the anti-statist case for a better welfare system more effective.

References

- Acosta-Ormaechea S., and Yoo, J. (2012) “Tax Composition and Growth: A Broad Cross-Country Perspective.” IMF Working Paper No. WP/12/257 (October). International Monetary Fund, Washington.
- Balcerowicz, L. (2015) “On the Economics and the Political Economy of Reforms.” *Decyzje* 24 (December): 67–90.
- Balcerowicz, L., and Rzońca, A., eds. (2015) *Puzzles of Economic Growth*. Washington: World Bank.
- Börsch-Supan, A. H. (2012) “Entitlement Reforms in Europe: Policy Mixes in the Current Pension Reform Process.” NBER Working Paper No. 18009 (April). National Bureau of Economic Research, Cambridge, Mass.
- Brooks, A. C. (2012) “America Already Is Europe.” *Wall Street Journal* (July 10).
- de Jasay, A. (2007) “Social Justice Examined with a Little Help from Adam Smith.” In J. C. Pardo and P. Schwartz (eds.) *Public*

- Choice and the Challenges of Democracy*. Cheltenham, U.K.: Edward Elgar.
- Feldstein, M. (2005) "Rethinking Social Insurance." *American Economic Review* 95 (1) (March): 1–24.
- Friedman, M. (1962) *Capitalism and Freedom*. Chicago: University of Chicago Press.
- Goulard, S. (2014) "France Remains in a State of Magical Thinking." *Financial Times* (October 1).
- Heutel, G. (2009) "Crowding Out and Crowding In of Private Donations and Government Grants." NBER Working Paper No. 15004 (May). National Bureau of Economic Research, Cambridge, Mass.
- Kelley, D. (1998) "The Emergence of Welfare Rights." In *A Life of One's Own: Individual Rights and the Welfare State*. Washington: Cato Institute.
- Lindbeck, A.; Nyberg, S.; and Weibull, J. W. (1999) "Social Norms and Economic Incentives in the Welfare State." *Quarterly Journal of Economics* 114 (1): 1–35.
- Morduch, J. (1999) "Between the State and the Market: Can Informal Insurance Patch the Safety Net?" *World Bank Research Observer* 14 (2): 187–207.
- Murray, C. A. (1984) *Losing Ground: American Social Policy, 1950–1980*. New York: Basic Books.
- _____ (2006) "A Plan to Replace the Welfare State." *Wall Street Journal* (March 22).
- Niskanen, W. A., (1996) "Welfare and Culture of Poverty." *Cato Journal* 16 (1): 1–15.
- _____ (2006) "Limiting Government: The Failure of Starve the Beast." *Cato Journal* 26 (3): 553–58.
- OECD (2014) *Economic Policy Reforms: Going for Growth*. Paris: OECD Publishing.
- _____ (2016) *Economic Policy Reforms: Going for Growth*. Paris: OECD Publishing.
- Tanner, M. D., and Hughes, C. (2015) "The Work versus Welfare Trade-Off: Europe." Cato Institute Policy Analysis No. 779 (August).
- Tocqueville, de, A. ([1835] 1997) *Memoir on Pauperism*. Chicago: Ivan R. Dee.
- Wicksell, K. (1967) "A New Principle of Just Taxation." In R. A. Musgrave and A. T. Peacock (eds.) *Classics in the Theory of Public Finance*. London: Macmillan.

WELFARE: SAVINGS NOT TAXATION

Roger Douglas and Robert MacCulloch

In many countries, the rising cost of publicly funded health care, retirement, and other welfare programs is forecast to put increasing pressure on government budgets. As a result, many governments are seeking to reform their welfare states so that costs to the taxpayer can be reduced, quality of outcomes increased, and the plight of low- and middle-income earners improved. Regrettably, there are currently at least three problems with much of the debate about reform of the welfare state.

First, disagreements are often focused around two opposing ideological viewpoints. One side is demanding more welfare spending and higher taxes, whereas the other is arguing for less welfare spending and lower taxes. Second, even when economists and others propose a welfare reform that appears promising in theory, designing the transition so that it is politically feasible is often overlooked. Third, the debates are typically quite narrow. They seldom focus on a comprehensive reform that would rewrite the rules governing the welfare and tax system as a whole, with the aim of making them work more fairly and efficiently.

This article shows how a country can move from a publicly funded welfare system to one that relies largely on private funding coming from compulsory savings accounts. The reforms we propose are designed to overcome the problems outlined above. We use New

Cato Journal, Vol. 38, No. 1 (Winter 2018). Copyright © Cato Institute. All rights reserved.

Sir Roger Douglas is head of Roger Douglas Associates and was Minister of Finance in New Zealand from 1984 to 1988. Robert MacCulloch is Matthew S. Abel Professor of Economics at the University of Auckland's Graduate School of Management.

Zealand, a nation with which we are familiar, as a case study, although the comprehensive reform we recommend can be adapted elsewhere.

How does it work? Taxes currently paid on personal earnings up to \$50,000 for single taxpayers go directly into the compulsory savings accounts.¹ A drop in the corporate tax rate and the removal of other government-imposed employee costs help employers fund contributions. These changes allow for *privately* funded welfare payments to substitute for public ones. Total spending levels can be maintained across most welfare categories, and transparent pricing of health care services and out-of-work cover can be introduced.

Provided that special privileges in the form of subsidies to businesses (i.e., corporate welfare) and grants to affluent families are discontinued, tax cuts can be made sufficiently deep to allow people to establish significant savings balances, while largely retaining pre-reform disposable incomes.

Even after our proposed tax cuts, the government retains sufficient revenues to act as an insurer of last resort, helping to pay for those individuals who cannot meet all of their own welfare expenses out of their savings accounts (e.g., the chronically ill).

Our “savings not taxes” reform offers the scope for efficiency gains, particularly in health care. While we believe these gains are plausible, they have not been factored into our estimated budgetary forecasts, which as a consequence probably understate the benefits of our reforms. One example of the scale of the possible gains comes from the experience of the pro-market reforms in New Zealand in the 1980s and 1990s (see Evans et al. 1996). Another example is Singapore, which uses compulsory savings accounts and currently spends just 4.8 percent of GDP on health and long-term care, compared with 17.2 percent in the United States and 9.5 percent in New Zealand, and yet maintains one of the highest quality services in the world.

More broadly, our “savings not taxation” reform is aimed at changing beliefs away from a culture of dependency to one of independence, whereby lower-income earners are given the opportunity to build up their own capital via tax relief and

¹All currency amounts in this article are in New Zealand dollars, unless otherwise stated (e.g., \$US).

employer contributions, and can then choose from among a range of affordable services.

Background: The Long-Run Viability of Publicly Funded Welfare

Large publicly funded welfare states are under threat. The dependency ratio, which is the proportion of elderly to younger, economically active workers, is expected to rise all over the world. Severe pressures will be brought to bear on pensions and, in particular, public health systems.

The ratio of public health and long-term care expenditure to GDP has already been rising steadily for several decades. The latest projections for the next several decades highlight the growing pressures. In the OECD's "cost-pressure" scenario, average health and long-term care public expenditures are projected to almost double, reaching approximately 14 percent of GDP by 2060. Furthermore, public pension spending is forecast to grow from 9.5 percent of GDP in 2015 to 11.7 percent of GDP in 2050 across OECD countries (OECD 2013).

Having recognized the welfare state problems that face almost every developed nation, we must ask ourselves, "What can we do about it?"

First, we need to quantify the problem in a way that politicians, economists, the news media, and others cannot ignore. One way is to forecast health and pensions spending as a fraction of GDP over the next several decades. As noted above, the OECD (2013) reports strongly rising trends. Another way is to measure the "fiscal gap," which is the present value of projected future government expenditures net of the present value of future taxes. Using this approach, Kotlikoff (2013) argues that the true U.S. fiscal debt is not the \$US 13 trillion usually reported by the government, but is instead over \$US 200 trillion.

Alternatively, one can bring the gross value of unfunded liabilities into the government accounts using accrual accounting principles.² New Zealand had a 2015 fiscal deficit of 4 percent of GDP once an allowance was made for the increased accrued

²The Public Finance Act 1989 in New Zealand promoted a move in this direction.

cost of health and pensions spending, which is due to the higher number of retirees at the end of the year compared with the beginning. That deficit stands in contrast to the cash surplus usually reported.

Second, we need to set aside traditional myths and return to fundamentals. This involves adopting several principles related to successful structural change (Douglas 1989):

- Only medium-term quality decisions, and not quick-fix solutions, make a difference. We must get the incentives and framework right to help ensure everyone acts more effectively.
- Decisions relating to welfare should identify and exploit economic and social linkages, so that every action will improve the working of the system as a whole. We should not treat every problem separately, as most countries do.
- Only large-scale reform packages provide the flexibility needed to demonstrate that any losses suffered by a group of people from one policy would be offset by gains for the same group in some other area.

Third, any fundamental reform of the welfare system would need to be based on improving opportunity, incentives, and choice if people are to accept widespread changes of the kind implemented in, for example, Singapore.

Singapore provides universal health care coverage at a lower cost than any other high-income nation. By most measures, such as infant mortality and life expectancy, outcomes are excellent. The cornerstone of the system is a compulsory medical savings account called MediSave. It is based on the idea that people should be helped to save for their own health care expenses. Individuals and their employers are required to contribute a specified portion of wages into each individual's account. The accounts are held within the government-managed Central Provident Fund.

Although MediSave funds belong to the contributing worker, the government has guidelines as to how the money can be spent. Its aim is to balance affordable health care against overconsumption and prevent the premature depletion of funds. For large bills that could otherwise drain one's MediSave funds, insurance schemes are available. The government offers a low-cost one called MediShield, under which individuals are automatically insured unless they choose to opt out.

Another component of Singapore's system is MediFund, which is a government safety net. It is a multibillion dollar endowment fund designed to help the lowest-income earners receive a level of care that they otherwise could not afford, even in the most highly subsidized public hospital wards.

Haseltine (2013: 51) believes that the MediSave account plays a crucial role in keeping high-quality health care affordable because,

when people have to spend their own money, as the Singapore system requires, they tend to be more economical in the solutions they pursue for their medical problems. In contrast, in countries with third-party reimbursement systems . . . since someone else is paying—government programs, insurance companies—there is little incentive to be prudent.

Our proposed reform includes the establishment of individual mandatory health savings accounts (HSAs), similar to the Singapore model.

Aside from the challenges that stem from maintaining a quality health care system, a large literature has discussed the challenges faced by public pension schemes due to population aging. When the number of recipients rises, governments often reduce the per capita generosity of publicly funded programs to help fit their budget constraints.

To prepare for this future, one might expect private savings rates to rise in nations with large publicly funded systems. However, this trend does not appear to be happening. Various ways of increasing savings in these nations are often debated (Shiller 2004). For example, some commentators advocate changes to the tax system to encourage savings. Others argue that there is a self-control problem biasing people toward overconsumption. In this case, automatically enrolling people in a savings plan, whereby one is joined up unless specifically electing to opt out, may be a solution. An example is New Zealand's Kiwi Saver scheme. Another response is to introduce compulsory retirement savings accounts, which are a feature of Singapore's system. While our proposed reform also introduces these types of accounts, it differs from the Singapore model by retaining the New Zealand state pension.

Last, we continue New Zealand's tradition of providing protection for the out of work. However, changes are made to the

existing system by establishing compulsory risk cover accounts to help pay expenses should one be unemployed, regardless of cause. In this sense, we depart from Singapore's zero unemployment benefit policy.

Designing the Shift to a "Savings Not Taxes" Welfare System

We now address the question of how to design a policy reform that allows a welfare system heavily reliant on public funding to be changed into one that increasingly draws on private funding. New Zealand is used as a case study, although we also discuss how the reform could be applied to the United States. A distinguishing feature of our new regime is that it proposes a unified approach to the funding of health care, retirement, and risk cover through the establishment of a set of compulsory savings accounts. In getting the framework right, we needed to adjust the tax system so that most New Zealanders of working age could provide for themselves.

Tax Reform

With respect to personal income tax (PIT) the rate is presently set at 10.5 percent for incomes from \$0 to \$14,000 and 17.5 percent for those between \$14,000 and \$48,000. Tax rates rise to 30 percent for incomes between \$48,000 and \$70,000. The top rate of PIT is 33 percent for incomes over \$70,000.

Under the new regime, PIT falls to zero for single taxpayers (i.e., a single person or couple with two incomes) earning less than \$50,000. It becomes 17.5 percent for incomes between \$50,000 and \$70,000, and 23 percent beyond \$70,000. For single-income families with dependent children, PIT rates fall to zero for incomes less than \$65,000. The corporate tax rate in New Zealand is cut from 28 to 17.5 cents per dollar of profit, and the goods and services tax (GST) rate is increased from 15 percent to 17.5 percent.

In total, taxes are cut by \$21.9 billion, or 9.1 percent of GDP (= \$239.5bn). This reduction is made up of a \$21 billion cut in personal taxes, a \$4.1 billion cut in company taxes, and a \$3.2 billion rise in GST.

Welfare Reform

At present, the New Zealand government funds welfare out of general taxation. In addition to health and risk cover (for unemployment, sickness, and disability) it also pays a pension. Under the new “savings not taxes” regime, the funds from the above tax cuts on income below \$50,000 (or \$65,000 for single-income families with children) go directly into the compulsory accounts. They are supplemented by an individual’s own contributions, and by the individual’s employer. Single taxpayers contribute 5 percent of earned income up to \$50,000. Their employers pay another 12.5 percent of income up to \$50,000. These add up to savings of \$17,500 per year for each person earning \$50,000 or more (and \$22,750 for a single-income family with children on \$65,000 or more).

The funds are used to help meet current health care and risk cover costs, as well as to build up savings for future payments in retirement. Smaller medical bills are paid directly whereas larger ones are funded by the purchase of catastrophic insurance plans. The government underwrites health and risk cover payments for those with insufficient savings (e.g., anyone who has been out of work for more than three years). Aggregate compulsory savings equal \$28 billion.

Health Care

The share of public health care spending in New Zealand has remained relatively constant over the last decade at around 80 percent of total spending (well above the average of 72 percent in OECD nations). In the OECD’s upside “cost-pressure” scenario, public health and long-term care spending is forecast to increase to 15.3 percent of GDP by 2060.

Our new regime changes the health system’s source of funding and introduces transparent pricing of health care services. Each person builds up an HSA that receives 45 percent of the total compulsory savings or \$7,875 ($= 45\% \times \$17,500$).³ Most of

³Cato Institute scholars first proposed (voluntary tax-incentivized) HSAs in the 1980s and were leaders in popularizing them among the public and policymakers. See, for example, Cannon (2006).

one's medical bills are paid out of this account. A prescribed level of savings is set for each person, and after it is reached the level of required savings is cut, increasing one's disposable income. Total contributions to the health accounts are \$12.6 billion ($= 45\% \times \28bn) each year.

An annual catastrophic health insurance policy must also be taken out to cover medical events costing more than \$20,000 in any one year (in 2015 dollars). This is paid for out of one's savings account. Those earning more than \$65,000 are expected to pay for part of their own health care, before drawing down their savings accounts. A 12.5 percent levy on the yearly health savings contributions of \$1.6 billion ($= 12.5\% \times \12.6bn) is made to help pay for the chronically ill, retired, and beneficiaries.

On the expenditure side, estimated drawdowns on the private HSAs in the first year of the reform equal \$7.5 billion. The government funds a further \$8.1 billion. As a result, the total amount spent on health remains the same at \$15.6 billion. In other words, the drop in public spending is fully offset by more spending from the compulsory accounts.

One would expect significant gains in service provision for the same level of funding. Efficiency gains arise from more transparent pricing of health care services, the ending of third-party funding, and the encouragement of more personal responsibility. As people directly spend their own money to pay for smaller health care bills and purchase their own insurance plans rather than having others do it on their behalf, behavior should change.

Haseltine (2013) suggests that the size of the potential efficiency gains may be enormous. For example, total public and private sector health care spending in Singapore, which uses a similar system of health savings accounts to those proposed here, is nearly one-quarter of the level in the United States. Yet Singapore's health outcomes are, if anything, better.

Retirement or Superannuation

At present, New Zealand pays a universal pension to people over 65 years of age who have completed modest residence requirements. It is flat rate (i.e., does not depend on previous income and is not means tested). The pension cost 5.1 percent of

GDP in 2015 and is forecast to rise to 8.1 percent by 2050.⁴ The Labour government also introduced the Kiwi Saver scheme in 2007, which is a voluntary retirement savings scheme. Employees are automatically enrolled and contribute a percentage of their gross earnings, unless they choose to opt out. Employers and the government also make contributions.

Under the new regime, individuals build up their own Superannuation Fund account, which receives 35 percent of their total compulsory savings or \$6,125 ($= 35\% \times \$17,500$). For many people these contributions will replace their existing Kiwi Saver payments. Our new regime extends the retirement age from 65 to 70 years old over the next 20 years (i.e., by three months per year) and retains the government pension, although its source of funding changes. At the start of the reform, the pension continues to be funded out of general taxation, though it will be increasingly covered by the increase in retirement age and a 25 percent tax levied on the size of each individual's Superannuation Fund on the date of retirement.

Total contributions to Superannuation Fund accounts are \$9.8 billion ($= 35\% \times \28bn) per year. Of this total, a pension levy of \$2.5 billion (in 2015 values) is paid on retirement. The remaining \$7.3 billion becomes savings that one is free to spend after retiring. In year one of the reform there are no withdrawals from accounts, whereas the government spends \$10.6 billion on pensions. As the age of retirement rises, public spending on pensions falls under the new regime, compared to the present one.

Risk Cover: Unemployment, Sickness, Invalid, and Accident Coverage

Unemployment benefits in New Zealand are currently paid out of general taxation and are unlimited in duration. The government also provides support for people with a health condition, injury, or disability, and sponsors the Accident Compensation Corporation, which

⁴Although the pension is paid out of general taxation, the Labour government established a Superannuation Fund in 2001 to help partially prefund future payments. Contributions are financed out of taxes but were suspended in 2009. Latest estimates suggest that about 8 percent of the expected cost of the pension in 2050 will come from the Superannuation Fund.

pays 80 percent of net wages to employed people who are unable to work due to accidents. These different schemes imply that the size of payments to an individual who is out of work depends on the cause of that unemployment. As a result there is an incentive to try to change category (e.g., from unemployed to sickness) in order to obtain a higher payment.

Under the new regime, individuals have a risk cover fund, which receives 20 percent of their total compulsory savings or \$3,500 ($= 20\% \times \$17,500$). A prescribed level of savings is set for the fund. Once reached, required contributions drop sharply. Should one become out of work, a drawdown occurs. If still out of work after 26 weeks, then a weekly payment is received from a catastrophic risk insurance policy (purchased with money in the fund). If one has insufficient funds in the savings account, or is jobless for more than 156 weeks, leaving one without insurance cover, then government assistance is given.

Total contributions to the risk accounts are \$5.6 billion ($= 20\% \times \28bn) per year. Estimated drawdowns in the first year of the reform equal \$1.5 billion and the government funds a further \$8.4 billion. Individuals have the choice to insure themselves at a higher level than the basic cover out of their own disposable income. Many higher-income earners would undoubtedly do so. Note that the payments that each out-of-work person receives no longer vary depending on the reason for absence from workforce.

Education

Primary and secondary schools in New Zealand are presently funded out of general taxation. Families may choose to pay for private education. The government also helps to fund early childhood education. There are no changes to the budget allocations for these programs under our proposed new regime, although an education tax credit would become available for any child whose family would like one.

University students, on the other hand, currently pay a subsidized fee for their degrees, and are also eligible for public grants. In 1992, the New Zealand government introduced a student loan scheme, which provides students with the opportunity to borrow for tuition fees, course costs, and living expenses. In 2006, student loans were made interest free.

Our reform retains the subsidized fee, but introduces a means test to restrict interest-free loans and grants to students from low-income, low-capital families. Note that university students may gain most from lower personal taxes under the new regime when they get older. The reduction in these grants equals \$3.3 billion. As a consequence, education funding falls to \$11.9 billion.

Other Welfare Expenses: Subsidies to Business, Kiwi Saver, and Working for Families

The government presently engages in a range of subsidies to business, which is sometimes called corporate welfare. It also subsidizes the Kiwi Saver scheme and funds the “Working for Families” (WFF) programme.

Corporate welfare includes a range of subsidies for broadband and fiber connections, movies that are “internationally focused and produced in NZ,” and support designed to maintain the strategic capabilities of industry. The total cost of these kinds of programs was \$1.35 billion in 2015. In addition, a range of accelerated depreciation tax allowances are available to businesses in the forestry, farming, bloodstock, and research industries, as well as favorable treatment of rental housing. These subsidies and allowances are discontinued under our new regime.

The WFF program, meanwhile, consists largely of earned income tax credits. Since our new regime cuts taxes for lower-income earners and helps people to establish their own savings accounts, both WFF tax credits and Kiwi Saver subsidies are no longer necessary. Even so, the WFF budget is still largely retained, but is instead reoriented to guarantee that the disposable income of low- and middle-income working families with dependent children doesn’t fall while their accounts are set up.

Of these different types of spending, there is a \$2.4 billion cut in subsidies to business and a drop in Kiwi Saver subsidies of \$720 million under the new regime (i.e., \$3.1 billion in total). As a result, “other” welfare spending falls to \$19 billion.

Impact on the New Zealand Government’s 2015–16 Budget

Under the present taxed-based system, revenues equal \$75.2 billion. The government spent the same amount of cash, mainly on the welfare state, representing 31.4 percent of GDP. The first major impact of the

new savings-based system is to cause tax revenues to fall to \$53.4 billion. This comprises lower personal income and company taxes, as well as lower interest and dividend taxes. These reductions are partly offset by a rise in GST revenues. A total of \$20.65 billion of the tax cuts is paid into the savings accounts for future welfare needs.

Consequently \$74 billion of funds are available for spending by the government and by private individuals on their own welfare needs under the new regime. Funds not spent out of individual accounts in the current year become savings for future expenditures. The impact on the New Zealand budgetary accounts for 2015–16 are summarized in Table 1.

Note that our regime gives the government a somewhat larger role in terms of its being the “insurer of last resort” compared with the current system in Singapore. In other words, we allow the state to provide more assistance to those who cannot fully fund their own catastrophic health insurance, or make payments for smaller bills, or both.

The second major impact of the new regime is the funding of personal savings accounts. A total of \$20.65 billion goes into these accounts to help contribute to each individual’s current and future personal health expenses and out-of-work income, and to pay for catastrophic health and risk cover insurance.

In the first year of the reform \$58.0 billion of welfare and other spending is paid by government, comprising \$8.1 billion on health, \$10.6 billion on pensions, \$8.4 billion on risk cover, \$11.9 billion on education, and \$19.0 billion on other expenses. By contrast, a total of \$8.8 billion of spending is funded out of the accounts, comprising \$7.5 billion on health and \$1.3 billion on risk cover.

A total of \$6.4 billion of subsidies to businesses and students from wealthy families are cut, which helps allow the tax cuts to be big enough to allow most people to establish significant savings balances. For a full set of current and forecast accounts, showing fiscal impacts, as well as details of the policy changes, see Douglas and MacCulloch (2016).

A Welfare “Savings Not Taxation” Reform for the United States

Although our focus has been on the New Zealand case study, the question arises as to whether this kind of reform would also be

WELFARE: SAVINGS NOT TAXATION

TABLE 1
NEW ZEALAND GOVERNMENT AND SAVINGS-BASED
BUDGETS FOR 2015–16

Row	(1) Government Budget (\$NZ millions)	(2) Savings-Based Budget (\$NZ millions)
Revenue Budget		
1		
Taxation (personal, corporate, goods and services tax)	75,200	53,350
2		
Funds paid into private savings accounts for current and future spending on health, risk cover, and retirement	0	20,650
Total Income	75,200	74,000
Expenditure Budget		
3		
Health, risk cover, retirement, education, and other government	68,800	58,000
4		
Health, risk cover, and retirement paid out of savings accounts	0	8,750
5		
Corporate welfare and grants to high-income earners	6,400	0
Total Expenditure	75,200	66,750
Government Cash Balance (= row 1 – row 3 – row 5)	0	(4,650)
Savings-Based Budget Balance (= row 2 – row 4)	0	11,900
Overall Balance	0	7,250

NOTE: The existing “taxes only” system is reported in column 1 and the effect of the “savings not taxation” system is reported in column 2.

SOURCE: New Zealand Treasury (2016): “Financial Statements of the Government of New Zealand for the Year Ended 30 June 2015.”

possible in the United States. Our answer is an emphatic “yes,” provided there is the political leadership and will to do it, and the required policies are packaged and presented in an appropriate way.

But why do we answer with a simple “yes”? Because the United States has so much room for reforming personal and business taxes. It would be relatively easy to lower tax rates dramatically, in return for the removal of various privileges. A reform of this nature would have many features that should appeal to both Republicans and Democrats. Since it replaces taxes and purchases of welfare services by third parties with mandatory savings that can be spent directly by individuals, the reform expands personal responsibility and freedom of choice. Furthermore, tax rate cuts are largest, percentage-wise, for low- and middle-income earners, and universal coverage is supported.

While we do not claim to be experts on the U.S. government’s revenue and expenditures, we are able to provide an outline of how the reform could be applied in this context.

Contributions to the personal savings accounts would be higher than in New Zealand, to cover higher health care costs. For an individual earning \$US 60,000, there would be a total of around \$US 21,000 in contributions each year. Of this total, the health care account would receive \$US 11,375, the retirement account \$US 6,125, and the risk account (for out-of-work coverage) \$US 3,500. As in the New Zealand version of the reform, annual catastrophic health and risk insurance policies would be taken out by each individual, with smaller bills paid directly from one’s mandatory savings accounts.

So how would these savings, equal to 35 cents on the dollar for someone earning \$US 60,000, be funded? It would look something like the following. Around \$US 10,500 comes by way of personal income tax cuts, \$US 3,000 from individual contributions, and \$US 7,500 from employers (totaling \$US 21,000). Although the government would lose \$US 10,500 in tax revenues (for our base-case individual), this loss is offset, in the main, by a cut in current and future health, retirement, and out-of-work public expenditures.

The individual contributions of \$US 3,000 would be paid in lieu of existing out-of-pocket health payments and retirement savings. As for the employer contributions of \$US 7,500, these are largely offset by reductions in the health, retirement, and out-of-work payments that they presently make. The reform would also include significant tax cuts for employers (including corporations).

Government retains certain important roles, however, both as “insurer of last resort” in the health care and out-of-work areas, and as provider of vital information (e.g., prices for health services). As an example of how transparency can be promoted, the Singaporean Health Ministry publishes prices on its website for medical conditions, surgeries, procedures, ward classes, and more. The aim is to empower people with information for making the best decisions regarding high-quality, low-cost care, and to encourage competition between providers.

We believe the United States has the opportunity to cut taxes well beyond what is necessary to fund the savings accounts, as set out above. One way to achieve this goal, while reducing the fiscal deficit, would be to cut government spending on huge privileges for special interest groups. These include grants to individuals—such as tax deductions on mortgage interest, low-income housing, business use of cars, and interest on student loans—as well as corporate welfare (i.e., subsidies to small businesses, large corporations, and industry organizations). Corporate welfare forms part of the programs of the Departments of Agriculture, Commerce, Energy, and Housing and Urban Development (De Haven 2012). Hundreds of billions of dollars could suddenly become available for personal and business tax cuts.

As noted earlier, the size of the potential efficiency gains from these changes is likely to be enormous. In health care alone, where the annual tax subsidy to employer-provided insurance has fueled the purchase of first dollar insurance plans, total public and private spending is presently 17.2 percent of GDP. However, Singapore’s experience suggests that, when individuals pay directly for smaller bills and buy catastrophic insurance cover out of their own savings accounts, health care spending can be reduced while outcomes improve. The welfare “savings not taxation” reform is designed to ensure that everyone has the funds to make these payments. It seeks to remove third-party payers, whether they are the government or employers.

In other words, the United States has a choice. It can maintain its complex tax code with all sorts of exemptions or, for example, have a tax system along the following lines: zero taxes on incomes from \$0 to \$US 60,000, a 20 percent rate from \$US 60,000 to \$US 120,000, and a 25 percent rate on incomes greater than \$US 120,000. People would, in these circumstances, become responsible for any additional

health care, retirement, or risk cover they wished to take out beyond what they received from their savings accounts.

In New Zealand, most corporate welfare was abolished in the 1980s and 1990s, based partly on a view that individuals are better investment managers than the state. Those changes also helped balance the government budget. For example, business concessions such as export incentives, accelerated depreciation allowances, and subsidies to agriculture were ended. Many individual concessions were also stopped, such as life insurance tax exemptions and rebates on interest payments for first home mortgages (Walker 1989).

The savings (and corresponding extra government income) that followed these changes allowed a key political strategy to be implemented in the form of large cuts in tax rates (e.g., the top personal income tax rate was halved from 66 percent to 33 percent). While corporate rates were also cut dramatically, the overall package proved to be politically popular, with the government of the day increasing the number of seats it held at the next election.

A similar package to the one outlined above, except this time for the United States, would prove to be equally popular, in our opinion. Significant efficiency gains are likely to arise from the changes to health care and risk cover, the removal of corporate welfare and other forms of privilege, and the tax cuts. Those changes would also allow individuals to accumulate large balances in their savings accounts to meet future health care and other needs in retirement. The fiscal gap of the U.S. government can be closed, and the system as a whole can become fully funded.

Conclusion

What kinds of outcomes can we expect from the new “savings not taxation” regime? First, consumers would now spend their own money (for health and risk cover) and save for their retirement. A culture of greater personal responsibility should develop, helping to control costs and increase quality. Consumers would become the principal buyers of welfare services, as they are for other goods and services, instead of third parties. Choice would become available for most people, not just the rich, creating a sense of empowerment.

Second, transparent price comparisons would become possible, particularly in the health services area, which should improve efficiency. Third, the primary role of government changes from funder

and provider to regulator, information provider, and insurer of last resort. Fourth, the personal goals of disadvantaged people would be recognized.

In summary, a reform of this type has the potential to lead to significant efficiency gains. It should also help secure the future of the welfare state, while at the same time retaining the necessary tax revenues to ensure universal coverage and equitable outcomes.

References

- Cannon, M. (2006) “Health Savings Accounts: Do the Critics have a Point?” Cato Institute Policy Analysis No. 569.
- De Haven, T. (2012) “Corporate Welfare in the Federal Budget.” Cato Institute Policy Analysis No. 703.
- Douglas, R. (1989) “The Politics of Successful Structural Reform.” Paper presented at the Mont Pelerin Society, Christchurch, New Zealand (November 28).
- Douglas, R., and MacCulloch, R. (2016) “Welfare: Savings Not Taxation.” Working Paper No. 31890, University of Auckland Economics Department.
- Evans, L.; Grimes, A.; Wilkinson, B.; and Teece, D. (1996) “Economic Reform in New Zealand 1984–95: The Pursuit of Efficiency.” *Journal of Economic Literature* 34 (4): 1856–1902.
- Haseltine, W. (2013) *Affordable Excellence: The Singapore Healthcare Story*. Washington: Brookings Institution.
- Kotlikoff, L. (2013) “Assessing Fiscal Sustainability.” Mercatus Center at George Mason University.
- New Zealand Treasury (2016) “New Zealand: Economic and Financial Overview 2016.”
- OECD (2013) “What Future for Health Spending?” OECD Economics Department Policy Notes No. 19 (June).
- Shiller, R. (2004) “Saving a World that Doesn’t Save.” Project Syndicate (February 6).
- Walker, S. (1989) *Rogernomics: Reshaping New Zealand’s Economy*. Auckland: New Zealand Centre for Independent Studies and GP Books.

LIQUIDITY RISK AFTER THE CRISIS

Allan M. Malz

The state of liquidity after the global financial crisis presents paradoxes. The money markets are awash in liquidity provided by central banks in most advanced countries. Yet by some indicators, liquidity appears to be impaired compared with precrisis conditions. By some market indicators, risk taking appears to be vibrant. By others, risk aversion and caution are dominant.

Also puzzling—and troubling—is a proliferation of market anomalies and oddities that have persisted or even grown more pronounced as the crisis itself recedes further into the past. These range from the failure of some basic near-arbitrage relationships to hold, to sporadic market “tantrums” and “flash crashes.”

Much criticism of postcrisis financial regulation has argued that it has made basic intermediary functions—lending and facilitating trades—costlier. The market paradoxes and anomalies indicate that something more is happening. They may be related to a more general impairment of market functioning, or to a rebuilding of precrisis levels of leverage, partly in hard-to-discern forms.

A reduction in liquidity compared with precrisis may not be unambiguously harmful. Liquidity is closely related to leverage, and ample liquidity before the crisis reflected the extensive leverage some market participants took on, enabled in large part by explicit or implicit public-sector repayment guarantees.

Cato Journal, Vol. 38, No. 1 (Winter 2018). Copyright © Cato Institute. All rights reserved.

Allan M. Malz teaches graduate courses in risk management at Columbia University. He has served as a risk manager at several financial firms and is a former Vice President at the Federal Reserve Bank of New York.

Liquidity is affected not only by regulations directly addressing liquidity, but also by policy measures affecting short-term funding markets, and the supply of and demand for liquidity in various forms. The impact on liquidity is not one to one with any particular change in policy. Frictions and tensions have been introduced that are hard to attribute to specific policies or trends. There is generally an identification problem in associating specific changes in the economy with specific policy changes. It is far more challenging when so many policy changes and changes in the overall economic environment have been taking place simultaneously.

But the results are potentially ominous. The still-highly leveraged financial system has a good chance of being subjected to unusual stresses soon as monetary policy normalizes in a low-growth environment and with high political risks. Will relatively modest reductions in central bank securities purchases or holdings be disruptive? Will the tensions be resolved “peacefully” or otherwise? The risk of disruptions emanating from surprising corners of the financial system is high.

What Is Liquidity Risk?

Liquidity is a difficult concept to pin down. It can refer to phenomena that seem quite disparate but are closely related: how easy it is to buy and sell assets, ease of access to credit, and the amount and forms of money in the financial system. Liquidity is intertwined with other phenomena, especially leverage and risk taking.

Market or transactions liquidity risk is the risk of moving an asset price against oneself while buying or selling, or covering a position, and thus being locked into it. It arises from the cost of searching for a counterparty and the cost of inducing someone else to hold a position. Market liquidity depends on participants’ positions, on institutions such as exchanges and dealer firms that aid search and make markets, and above all on the risk appetites of market participants.¹

Funding liquidity risk is the risk to market participants of being unable to maintain debt financing, and having as a result to liquidate a position at a loss that they otherwise would keep. Funding liquidity risk events typically involve short-term debt, which rolls over more frequently than long-term. Financial firms are most susceptible, since they are often in the business of using short-term debt to finance longer-term financial assets, which are harder to sell

¹Foucault, Pagano, and Roell (2013) is a general introduction to market liquidity risk.

without loss. The yield curve is generally upward-sloping, making maturity mismatches attractive.

Funding liquidity depends on both the reality and perceptions of borrowers' creditworthiness and is more fragile when borrowers are highly leveraged. This fragility can be even greater when there is uncertainty about how leveraged the borrowers are; large banks' balance sheets, for example, are notoriously opaque. Short-term credit is granted in part based on public-sector guarantees, adding to fragility.

Liquidity is closely connected to risk appetites. Higher risk aversion renders liquidity fragile; perceptions about creditworthiness or the solidity of guarantees may suddenly seem less well-founded. When risk aversion is high, dealers are more reluctant to make markets in size and lenders, including short-term lenders on collateral, are less eager to extend credit and to put up with sparse information on creditworthiness.

Funding liquidity is also closely linked to market liquidity, since they interact, and since both suffer when risk aversion is high. Banks, dealers, and traders depend on funding liquidity, either to finance inventories or provide collateral, in order to make markets. Short-term lenders depend on market liquidity to support borrowers' ability to repay and the value of the collateral they hold. If many borrowers experience a liquidity risk event at the same time, or market liquidity for the assets being financed is poor, forced unwinding or positions may result in a "fire sale," with prices driven far from any notion of fair value.

Financial crises are often triggered by liquidity events coinciding with abrupt changes in sentiment. Liquidity risk events in financial markets are apt to affect many market participants at the same time. The withdrawal of short-term lenders may be effected through higher interest rates, but as often through nonprice rationing, such as increases in "haircuts" on collateral, refusal to accept some collateral, or simple refusal to lend. The onset of a crisis is often described as a liquidity crunch, but could as easily be described as a sudden spike in risk aversion, expressed in part as aversion to some types of privately created liquidity. It is often marked by a run or run-like behavior, in which short-term lenders suddenly converge on a borrower.²

²Shleifer and Vishny (2011) is an overview and Acharya and Schnabl (2010) a case study of the role of funding liquidity risk in crises and its relation to market liquidity risk.

The financial system creates most of what is used as money, in the form of assets corresponding to some short-term debt. Most of the narrowly defined money supply consists of bank deposits. Other short-term assets included in standard definitions of monetary aggregates also carry out at least some of the functions of money, such as availability at par or at least a highly predictable value to buy other assets. A large volume of assets with these characteristics, but not necessarily included in monetary aggregates as defined by central banks, is created by financial intermediaries in wholesale short-term funding markets, in which larger financial and nonfinancial firms participate. Also known as shadow banking, they have grown rapidly over the past four decades.

The spread between short-term rates with terms of a few days or weeks and longer money market rates is generally particularly wide. Part of the term spread at the very short end of the yield curve is a money premium, the interest forgone in exchange for money services. The money premium makes short-term wholesale funding even more attractive (Greenwood, Hanson, and Stein 2015).

Two important mechanisms of private money creation outside banking are repurchase agreement or repo markets and money market mutual funds (MMMFs). Repo is a form of short-term collateralized lending in which the lender of cash receives bonds as collateral from the borrower. Repo is legally framed as a pair of bond purchases and sales, one now and the other in the future, both at prices that are determined now and imply a lending rate. Repo adds to the supply of liquidity because holders of repo-eligible securities can use them to quickly obtain a predictable amount of liquidity at a predictable rate. It has become a large part of short-term wholesale lending, particularly since a regulatory change in the 1980s exempting repo from the automatic stay in bankruptcy, thus permitting repo cash lenders to immediately sell collateral if their counterparty defaults.³

MMMFs first arose in 1971 and grew rapidly from the mid-1970s, a time of high and rising interest rates, as a means of evading the regulatory ceiling on bank deposit rates. From 1983, the Securities and Exchange Commission (SEC) set credit standards for MMMFs' asset pools and permitted them to use a fixed \$1 par value for share transactions. Growth of so-called institutional MMMFs serving

³Baklanova, Copeland, and McCaughrin (2015) is an introduction to repo.

nonfinancial firms, used less as ready cash than as liquidity reserves, was even faster than that of retail MMMFs. MMMF assets reached nearly \$4 trillion before the crisis, the bulk initially in commercial paper, but displaced in part over time by short-term municipal and corporate bonds, and increasingly, repo.

Policy Changes and Interest Rates Since the Crisis

Market behavior has been affected by major changes in monetary and regulatory policies. Policy has in some respects added tremendously to both the supply of and demand for liquidity. Liquidity is also affected by nonpolicy trends that began before the crisis, but have been transformed in its wake, mostly associated with low interest rates. We'll summarize these developments to help understand recent market behavior.

The identification problem makes it difficult to say which policy or economic change has led to which impact on liquidity. But, in some cases, some direction of influence can be stated.⁴

Monetary and Debt Management Policies

The Federal Reserve initially responded to the crisis with conventional monetary easing and emergency programs that combined liquidity support for the financial system as a whole with targeted credit support for specific firms and types of intermediation. For most of the past decade, the main policy tool in the major industrial countries has been variants of quantitative easing, the purchase of bonds in large volumes on the open market. It is intended to further reduce the long-term rates most relevant for investment by reducing term and risk premiums once short-term rates have been brought near zero.⁵

Forward guidance through public statements, on which the Fed has placed far greater reliance since the crisis, is intended to support easing by committing the Fed to keep short-term rates near zero in the future and thus lowering the expected future rates component of the long-term rate. Forward guidance adds to

⁴The identification problem also makes it difficult to estimate the compliance cost of new and revised regulation, which often leads to different ways of conducting business, not just additional staff and systems.

⁵Ihrig, Meade, and Weinbach (2015) is an overview of the Fed's new policy tools and its choices for an exit from crisis policies.

liquidity in the same way that interest-rate smoothing did prior to the crisis, by making market participants confident that financing rates will stay steady.

Quantitative easing is intended to work through different channels from conventional monetary policy, which relies on managing overnight liquidity to set the overnight federal funds rate, focusing on central bank assets rather than liabilities. But by creating a large volume of central bank reserves, it has a strong impact on the markets in which the funds rate is determined. The most direct impact is to flood markets with safe overnight liquidity and vastly reduce banks' need to trade them.

Several new policy tools have been introduced in the United States to manage that impact and preserve the fed funds market for a future return to conventional policy. Interest on reserves was introduced in late 2008 to keep the funds rate near the target rate or within the target range once the Fed permitted its balance sheet to expand to accommodate the emergency programs. Overnight reverse repurchase agreements (ON RRP) were introduced as a floor on the funds rate once the bottom of the target range was raised above zero. Policymakers have expressed concern that, by providing a flight destination from private forms of liquidity, ON RRP could be destabilizing in times of financial stress. While lower limits have at times been placed on ON RRP volume, it is currently limited only by the size of the Fed's portfolio of Treasury bonds.

U.S. Treasury cash management practices were adapted early in the crisis to support the Fed's efforts, and more recently have been an independent factor influencing markets. All these changes have affected all money markets, not only that for fed funds.

Regulatory Policies

Regulatory policies have changed substantially since the crisis. Many new rules, a number ordained by the 2010 Dodd-Frank Act, have been adopted in the United States, though the compliance deadlines for some have not yet arrived.⁶

⁶See Aaron, Demers, and Durr (2015) for an overview of the leverage and liquidity rules and the likely direction of their impact. The Basel capital and liquidity standards are stated in Basel Committee on Banking Supervision (2011, 2013), among other standards documents.

The Basel minimum capital standards have been revised and increased in a several ways. The minimum common equity ratio is now higher. Criteria for the recognition of nonequity capital have been tightened, but most of the additional capital required can be funded in nonequity forms. Dividends and other payouts are to be slowed or halted if additional capital buffers are breached. Some apply to the largest banks, including a requirement to issue non-equity capital that is intended to be zeroed out and thus limit the need to apply public funds in the event of insolvency.

The revised Basel and U.S. capital standards retain the precrisis reliance on risk-weighted assets, and the minimum ratios remain low compared with previous eras; but a capital ratio based on the size of the bank's balance sheet has been introduced. Banks must satisfy the higher of the two, so either can become a binding constraint. The size-based capital ratios also have a more stringent version for certain very large U.S. bank holding companies, insured bank subsidiaries of which are subject to an enhanced Supplementary Leverage Ratio (SLR) of 6 percent.

Leverage ratios endeavor to address two drawbacks of risk-weighted ratios. Whether based on a table of weights supplied by the regulator or on internal models, weights designed to be risk-sensitive may be drastically inaccurate and induce banks to shift to higher-risk assets. Two classic precrisis examples are the zero or near-zero weights on sovereign debt and on senior securitization tranches with low-quality underlying asset pools. The second drawback is that risk-weighting systems, if they are to be accurate and comprehensive, must also be complex in design, on top of the complexities of the composition of regulatory capital. Together with the presence of public-sector guarantees, these problems enable banks to take risk beyond what their capital funding can appropriately bear without external guarantees.

The advantages of a regulatory leverage ratio in countering these problems, however, have the inverse consequence of disincentivizing transparent low-risk activities that "use a lot of balance sheet," that is, add large positions counted as assets in the denominator of the ratio. The U.S. deposit insurance program now imposes fees based on total assets rather than insured deposits, with a similar effect. Criteria for inclusion or exclusion of on- and off-balance sheet positions in the denominator are themselves complex, in addition to the complexities of the capital numerator.

One of the goals of postcrisis regulatory reform has been to reduce short-term wholesale funding. Higher capital standards work in that direction, but new liquidity regulations have also been introduced. Banks' thin capital cushions and funding liquidity risk are treated as requiring distinct regulatory approaches, rather than viewing short-term wholesale funding as a form of leverage that is particularly attractive to large banks with stronger implicit guarantees.

To that end, two new minimum ratios have been introduced. The Liquidity Coverage Ratio (LCR) can be thought of as a 30-day liquidity stress test, requiring banks to keep enough liquid assets to meet a specific scenario of a run on their short-term funding. The Net Stable Funding Ratio (NSFR) limits maturity mismatch on banks' balance sheets. Both rely on complex rules for inclusion, weighting and exclusion of assets, liabilities, and off-balance sheet items. The ratios penalize short-term funding not secured by low-risk collateral (i.e., government bonds) and reward short-term lending, and would *ceteris paribus* increase the money premium.

There is a tension between the liquidity and capital standards as they now stand. Deposits with central banks—primarily excess reserves in the United States—are near-riskless. But they lengthen the balance sheet and raise minimum capital if the SLR is binding and are thus costly to finance relative to their risk and return. But these excess reserves can be used to satisfy the new liquidity regulations, so banks reap a regulatory compliance reward, in addition to interest on reserves.

The so-called Volcker Rule prohibits proprietary trading, as well as ownership by banks, which benefit from deposit insurance and lender of last resort support, of hedge funds, private equity funds, and loan securitizations. Because of the difficulty of discriminating between positions held in anticipation of price changes and those held to facilitate market-making, hedging, and other permitted activities, the Volcker Rule is accompanied by extensive recordkeeping requirements, and its implementation will depend on regulators' interpretation even more than for other postcrisis rules.

Changes in the rules affecting MMMFs include permitting a fixed net asset value (NAV) only for funds that invest in U.S. government-issued securities and repo, or serve exclusively retail investors. Other MMMF types must mark their NAV to market daily. Moreover, if there are substantial redemptions in a short time from

a nongovernment fund, it is obliged to restrict or impose fees on further redemptions.

The revised rules shift demand from corporate, and to a lesser extent municipal securities, to U.S. issues. The share of government-only funds in total MMMF assets has grown from about one-third in mid-2015 to about three-quarters. The share of U.S. government securities and repo in the assets of all MMMFs has also risen. The shift in allocation enlarged the market for the Fed's floor-setting ON RRP.⁷

Some changes have been introduced through persuasion more than regulation. Repo trading is carried out in different trading venues. Some repo trades are focused on financing long or short positions in specific bonds used as collateral. These repo trades between dealers and their customers are usually cleared bilaterally, a process called delivery versus payment. Dealers themselves want to have matched books, that is, facilitate trades for customers, but limit their own net risk, and interdealer trading that achieves this has been moving to a facility called the General Collateral Financing (GCF) Repo Service. Repo trading focused on lending or borrowing money, rather than on specific securities, including GCF repo, settles on the triparty repo platform, a service that has been provided by two large U.S. banks, JP Morgan and BNY Mellon.

A major supervisory concern has been the extremely large exposure of the two triparty clearing banks during regular business hours. The banks, one of which is exiting the business, fund repo dealers between the time expiring contracts roll off and the dealers return cash to their counterparties in the morning, and the time they enter into new contracts in the afternoon. The Federal Reserve has been pressuring the banks to reduce this intraday exposure and dealers to use GCF repo. The intraday exposure has in fact been greatly reduced. The effort is consistent with the regulatory effort to channel derivatives trading from bilateral to multilateral clearing platforms.

A relatively new and still diffuse regulatory effort addresses financial stability through a so-called macroprudential approach to regulation, which views regulation from the point of view of the

⁷The shift in demand was also reflected in a widening of spreads between commercial paper and Treasury bills, and of term spreads in the unsecured interbank lending market, which have since been reversed.

stability of the financial system as a whole, in addition to the traditional microprudential focus on the safety and soundness of individual financial firms. The approach focuses on externalities and systemic risk in finance, that is, profitable activities of individual intermediaries that are said to impose uncompensated risks on others. While not yet well articulated, it has motivated liquidity regulation, as well as supervisory guidance restrictions on riskier loans by banks.⁸

The macroprudential approach creates tensions with monetary policy. The approach could aim to crack down on symptoms of excessive credit expansion through regulatory measures. This would insulate monetary policy from financial stability considerations and permit longer periods of ease. Or it might add countercyclical regulatory measures to a generally tight monetary stance that is wary of credit expansion. In the past, this debate was known as “lean or clean.” Today, it is related to the debate on the pace of normalization in the absence of rising inflation.⁹

Low Interest Rates and Demand for Safe Assets

Some precrisis trends have continued since the crisis. Interest rates, particularly long-term rates, had fallen steadily in the years leading up to the crisis. The real interest rate is not directly observable, econometric estimates have wide confidence bands, and market-based estimates are bundled with unobservable risk premiums. But the estimates agree that the real rate has been falling steadily, worldwide. The U.S. real rate has fallen about an additional 200 basis points since the onset of the global financial crisis.

The demand for “safe assets,” including highly rated sovereign bonds and the short-term assets resulting from private liquidity creation described earlier, is closely related to low rates. Before the crisis, the two phenomena met in the “conundrum,” the unusually

⁸Kim, Plosser, and Santos (2017) describe some of the problems arising from reliance on guidance in supervision.

⁹An inverse argument is presented by Greenwood, Hanson, and Stein (2016), who propose the Fed maintain a large balance sheet for the foreseeable future. Its liabilities would substitute for private-sector money-like assets that are vulnerable to runs. The authors argue this would be a better-aimed method for averting such runs than the SLR. They address the concern about flight into Fed liabilities by limiting expansion during stress periods.

flat term structure of interest rates, reflecting high demand for longer-term bonds. It was attributed to demand for low-risk fixed-income assets by an aging, wealthier world population. More recently, attention has focused on the ability of such assets to serve liquidity functions, as described above. The crisis is said to have eliminated some asset classes from the supply of safe assets by exposing their credit risk, adding to the downward pressure on interest rates.¹⁰

The demand for safe assets contributes to strong demand for U.S. dollar-denominated assets. The demand for U.S. dollar intermediation, including high volumes of both dollar funding and dollar-denominated assets, was seen prior to the crisis and continues after. Much dollar funding is raised directly by non-U.S. issuers of dollar-denominated bonds, and much of that in recent years has been in emerging markets. But a great deal of dollar funding is raised by U.S. offices of foreign banks (FBOs) and lent both within and outside the United States.

Is Liquidity Impaired Now?

Financial markets are displaying many unusual behaviors: conflicting indicators of market and funding liquidity, anomalies and impairments of normal market functioning, and conflicting indicators of risk appetite. These behaviors are related to one another, and how the conflicting evidence is resolved will influence how well the economy can respond to shocks and how disruptive will be central bank efforts to return to more normal monetary policies.

Market liquidity, the ability to alter holdings quickly and cheaply, is very hard to measure. Observable phenomena such as trading volumes and bid-ask spreads are only indirectly related to the cost of buying and selling, and to how much a large securities order changes the market price, and for how long. Even data on observable characteristics are hard to obtain and summarize across asset markets.

Concerns have been raised that higher capital and new liquidity requirements and the Volcker Rule have inhibited market liquidity

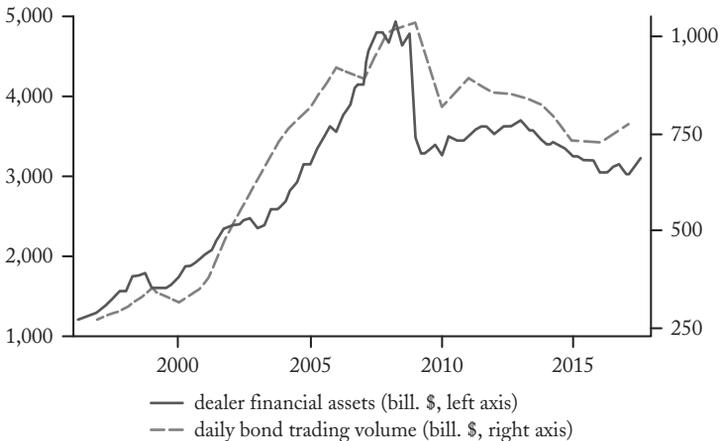
¹⁰Caballero, Farhi, and Gourinchas (2017) is a recent survey. The authors take the view that the effective lower bound in interest rates can prevent the market for safe assets from clearing, leading to elevated risk premiums and forcing the market to clear via a reduction in economic activity.

since the crisis, particularly in the U.S. corporate bond market. But the evidence isn't robust. Market structure has also been changing, for example the growing prevalence of electronic trading, so identification of the cause of change is difficult.

Trading volume of U.S. bonds overall and dealer balance sheets have declined since the crisis (Figure 1), although trading in corporate bonds, a small fraction of the total, has increased a bit. Much of that may be due to the increased issuance of corporate bonds. Turnover of corporate bonds, the frequency with which the bonds are traded, appears to have declined over the postcrisis decade.

Bid-ask spreads, however, don't appear to have greatly changed. That would be consistent with the shrinking of dealer balance sheets and a shift from market makers acting as principals and taking long and short positions, to agency trading. It would also be consistent with a shift away from low-risk intermediation requiring balance sheet, and thus debt or equity funding. But entering or exiting a position may now take longer and require more, smaller transactions. All in all, the upfront cost of trading may be steady or even

FIGURE 1
U.S. DEALER ASSETS AND BOND MARKET
TRADING VOLUME, 1996–2017



SOURCES: Total financial assets of security brokers and dealers, \$bill. Federal Reserve Board, Financial Accounts of the United States (Z.1), Table L.130. Average daily trading volume of U.S. bonds, \$bill. Securities Industry and Financial Markets Association (SIFMA).

declining, but less immediacy in trades imposes additional implicit costs.¹¹

Since the crisis, transactions liquidity in specific markets has suddenly become not just impaired but has virtually disappeared for periods of a few minutes to a few days. These episodes are loosely referred to as “tantrums” and “flash crashes,” and have occurred in a range of markets and market types. An early such episode took place on May 6, 2010, in the U.S. equity market. Others have occurred in U.S. Treasury and German Bund markets, as well as in the sterling markets. A recent episode in the silver markets is typical of the intra-day variety. On July 6, 2017, silver futures fell about 10 percent early in the Pacific trading day, and recovered within a few minutes.

Market as well as funding liquidity may appear ample if demand to use it is low, but disappears if uncertainty and risk aversion abruptly rise. The experience during tantrums suggests that market liquidity may become much less reliable if there is a sustained change in sentiment from the current low volatility regime. The concern about liquidity in a stressed market is strengthened by evidence from the U.S. corporate bond market that market liquidity of specific issues is badly affected by downgrades.¹² The tantrums also suggest that market liquidity can’t be assessed on its own but only in the context of the other unusual market behaviors we’re describing.

Risk Appetite

Evidence on risk appetites is also in conflict, with evidence of both risk seeking and risk avoidance. Concerns had been raised prior to the crisis about “reaching for yield,” the propensity of unleveraged long-term investors such as pension funds and insurance companies with fixed future retirement or claim liabilities to shift allocations to higher-risk assets to compensate for low interest rates. It may also be driven by regulatory constraints and public guarantees enjoyed by some intermediaries. For example, regulatory capital requirements on U.S. and European insurers are lower if they hold bonds rather

¹¹See, for example, Mizrach (2015), Adrian et al. (2017), and International Organization of Securities Commissions (2017). Anecdotal evidence from portfolio managers, as well as a report on European corporate bond market liquidity (European Commission 2017), is decidedly more negative on the state of corporate bond market liquidity.

¹² See Bao, O’Hara, and Zhou (2016).

than equities as investments, and regulatory and accounting rules constrain pension plans to hold bonds rather than equities against some liabilities. But within those constraints, long-term investors may increase the duration or credit risk of loan and bond portfolios, select the riskiest bonds within a regulatory category, or deploy their equity allocations in riskier forms such as hedge funds and private investments. Bond investors, before the crisis, pressured rating agencies to inflate ratings by assigning a higher fraction of securitization pools to the most senior bonds (Calomiris 2009).

The precrisis patterns of reaching for yield and demand for safe assets have continued barely interrupted after the crisis. For example, the allocation to lower-quality bonds (BBB-rated and speculative grade) in U.S. property and casualty insurance fixed income sector portfolios has risen from about 7 percent in 2007 to about 21 percent in 2016 (New England Asset Management 2017).

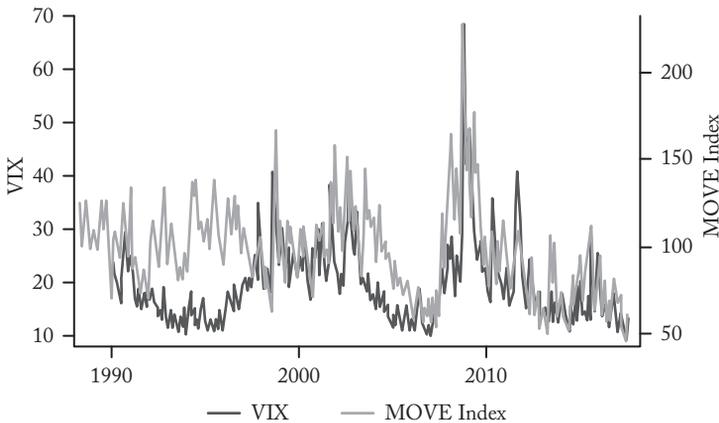
Implied and realized asset return volatility have been extremely low in recent years. Implied volatility is expressed through option prices, and it is low when market participants expect future realized volatility to be low, but also when they are eager to take the risk of selling puts and reluctant to pay for the corresponding protection, even at relatively low prices. Figure 2 illustrates with the VIX, a well-known index of equity implied volatility, and the MOVE index of U.S. interest rate volatility. Both are at the lowest levels of their quarter-century history.

Implied volatility is generally close to, but somewhat higher than, expected volatility, measured using recent realized volatility. The difference, or variance risk premium, is a measure of the pure reward to supplying protection against volatility, and has been substantially lower in the post- than in the precrisis years.¹³

Like low interest rates, low volatility is a precrisis phenomenon that has reached new extremes. It is related to reaching for yield, as both express an increased supply of implicit liquidity puts at given prices. Stress conditions have a similar impact on option sellers and investors

¹³See Bollerslev, Tauchen, and Zhou (2009) and subsequent literature. A simple measure of the variance premium is the difference between the VIX and an exponentially weighted moving average measure of S&P 500 index return volatility (using a decay factor of 0.94). The average daily difference in annualized volatilities from January 1990 to the onset of the crisis at the end of February 2007 was about 4.25 percentage points. The average difference from late September 2011, when the “hot” phase of the crisis can be said to have calmed down, to present was about 2.75 percentage points.

FIGURE 2
EQUITY AND BOND IMPLIED VOLATILITY, 1988–2017



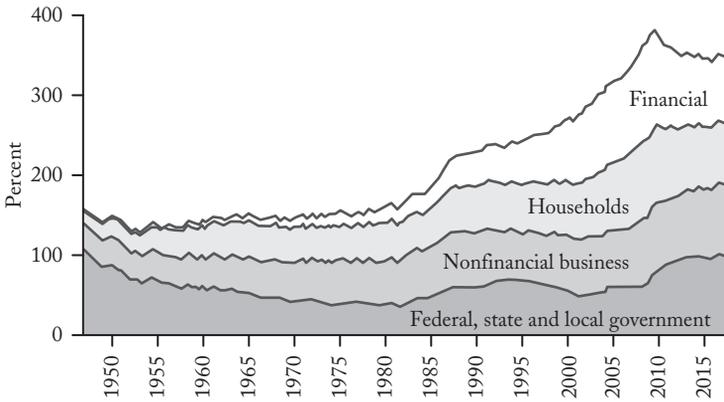
SOURCES: CBOE Volatility Index (VIX), weighted average of prices of options on the S&P 500 Index with approximately one month to expiry and with a range of strike prices. Merrill Lynch Option Volatility Estimate (MOVE) Index of implied normal volatilities of at-the-money options with approximately one month to expiry on U.S. Treasury notes and bonds. Bloomberg LP.

in down-in-credit bonds: both find themselves locked into positions that can only be exited at prices that have just changed sharply for the worse. Both volatility selling and reaching for yield can lead to discontinuous, “tantrum” behavior of markets.

Leverage, finally, remains very high, just below precrisis levels. The financial sector has reduced debt levels, and households modestly, while the nonfinancial business and public sectors have increased borrowing (Figure 3). But the rapid increase in debt levels of the post-stagflation era hasn’t been reversed.

Yet there is also countervailing evidence on risk appetites. One indicator of high risk aversion is the same low real interest rates said to motivate reaching for yield. The flat yield curve, the strong demand for 100-year bonds, and the implied volatilities discussed above suggest strong expectations that nominal rates will remain low for the foreseeable future. Some of the factors that could be causing this decline are consistent with or point to risk aversion: a decline in growth and anticipated future real returns, and higher demand for safe assets.

FIGURE 3
U.S. DEBT-TO-GDP RATIO BY SECTOR, 1946–2017



SOURCES: Ratio of total debt outstanding (debt securities and loans) to GDP, current dollars, percent, quarterly, Q4 1946 to Q2 2017. Federal Reserve Board, Financial Accounts of the United States (Z.1), Tables D.3 and F.2.

Lower prospective growth and returns are consistent with a range of other observations. Rates of capital investment, the growth of labor productivity, new business formation, and growth of bank loans and leases, for example, are lower than before the crisis. There are alternative explanations for them (e.g., the shift to a service economy, measurement issues, and the role of technology), but these data are also quite consistent with low prospective returns.¹⁴

The demand for safe assets may also have increased in recent decades due to factors other than risk aversion, such as increased wealth and aging populations in both developed and developing countries. Some observers have also argued that the supply of safe assets has declined due to central bank purchases and the

¹⁴Two examples: The entry rate of new establishments—places of business—in the United States fell from an average of 12 percent in the decade before the crisis to 10 percent since the crisis, at times below the exit rate (U.S. Census Bureau, Business Dynamics Statistics). Loans and leases in bank credit have grown about 5 percent annually over the past 5 years, a period starting 3 years after the last NBER business cycle trough, compared to about 8 percent over the entire postwar era and over the 20 years prior to 2007 (Federal Reserve H.8 release).

disqualification of securitized products considered low-risk prior to the crisis. The effect on collateral supply of central bank purchases is only partially offset by adding short-term safe assets in the form of reserves.

The persistence of risk aversion can be seen in corporate bond spreads over risk-free rates, which have tightened massively since the crisis, but are not at record lows. Similarly, Libor remains at a wider spread to overnight interest rate swap rates than before the crisis. Money-market credit spreads may also be kept wider by the combined impact of the LCR and MMMF reform, which has made banks less eager to receive deposits of corporate cash balances not intended for daily transactions, pushing them to government-only funds, and has made government paper relatively attractive. Issuance of commercial paper by U.S. banks other than FBOs has also declined.¹⁵

Risk aversion can be seen even in equity prices. Although some commonly cited measures, such as the Shiller CAPE, are high, they are in part an artifact of low interest rates.¹⁶ An alternative measure takes low rates into account by subtracting an estimate of the real interest rate from the current dividend yield. Both are real, rather than nominal rates, and the spread is a simple measure of the real equity excess return or risk premium, in effect treating equity as a bond that has both credit and cash-flow risk. The dividend-real yield spread is very high by historical standards, higher even than during the bear market of the late 1970s (Figure 4).¹⁷

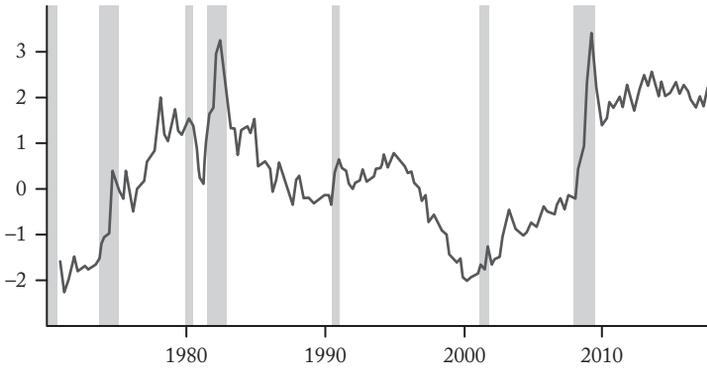
But low interest rates may not be due only to low growth and pessimism about future returns. Persistent and surprisingly low inflation and inflation expectations are a large part of surprisingly low nominal rates. There is wide agreement that low inflation is a result of past

¹⁵The lows in the BofA Merrill Lynch U.S. Corporate Master Option-Adjusted Spread (OAS) data are in 1997, and a substantially higher precrisis low was reached in March 2005. As of early October 2017, OAS for all ratings, as well as the BBB-AAA quality spread, were higher than their March 11, 2005, levels. Astonishingly, even in the eurozone, where the quantitative easing policy has included large-scale purchases of corporate bonds, the Bloomberg Barclays Euro-Aggregate Corporates OAS is three times as high as in March 2005. The BofA Merrill Lynch Euro High Yield Index OAS is also materially higher than its mid-2007 low.

¹⁶The Shiller CAPE is currently over 30, its highest value since the decline in 2000–01.

¹⁷This observation is another way of seeing the safe asset shortage.

FIGURE 4
DIVIDEND–REAL RATE YIELD SPREAD, 1970–2017



SOURCES: Twelve-month trailing dividend yield of the S&P 500 index (Bloomberg LP) minus Laubach-Williams estimate of the short-term natural rate (available at http://www.frbsf.org/economic-research/files/Laubach_Williams_updated_estimates.xlsx), percent, Q4 1970–Q2 2017.

success in anchoring expectation inflations and current low growth rates.¹⁸ But while different approaches to the relationship between low inflation and real rates point in the same direction for monetary policy, they point in opposite directions for an understanding of risk appetites.

In one view, the market rate is generally near the equilibrium real rate. Inflation is suppressed because both the real and nominal rates are near zero. Rather than inflation expectations dictating the gap between real and nominal rates, the low nominal rate, together with a real rate immune from longer-term policy influence, dictates low inflation expectations. In this view, pessimism and risk aversion are ascendant. In another view, and despite the evidence to the contrary provided by dormant inflation, the current market rate may be below the real rate. The relatively low market rate may then be an incentive to risk taking. Both views imply a similar policy conclusion: raise nominal rates. But the motivations are quite different: on the one

¹⁸ Apart from survey data and yields on inflation-indexed U.S. Treasury securities, the anchoring of inflation expectations can be seen in the strongly positive correlation between stock market returns and changes in government bond yields. The correlation had been negative during the 1970s, turned positive around 2000, and increased sharply during the crisis.

hand, to provide room for inflation to rise, and thus permit at least a transitory rise in growth, and, on the other, to equalize the market to the real rate and tamp down financial excesses.¹⁹ The debate is related to the paradox of volatility, the likelihood that periods of unusually buoyant risk appetite, ample liquidity, and low volatility tend toward instability.

Arbitrage Failures and Other Anomalies

Many anomalies have appeared, primarily in the money markets, in which near-riskless arbitrage opportunities fail to disappear over time. Banks are apparently unwilling to lend at a virtually risk-free spread in these cases. These anomalies persist either because banks and other intermediaries aren't devoting sufficient capital to exploiting them, or because of market frictions that are new or worse since the crisis.

Persistent arbitrage opportunities are not a new phenomenon. The arbitrage process is almost never entirely free of risk, and takes place slowly and incompletely due to various market frictions.²⁰ Some arbitrage opportunities are astonishingly large and persistent, and preceded the crisis, such as the difference between estimates of future inflation rates expressed in nominal and indexed U.S. Treasury securities, and fixed rates of inflation swaps (see Fleckenstein, Longstaff, and Lustig 2014).

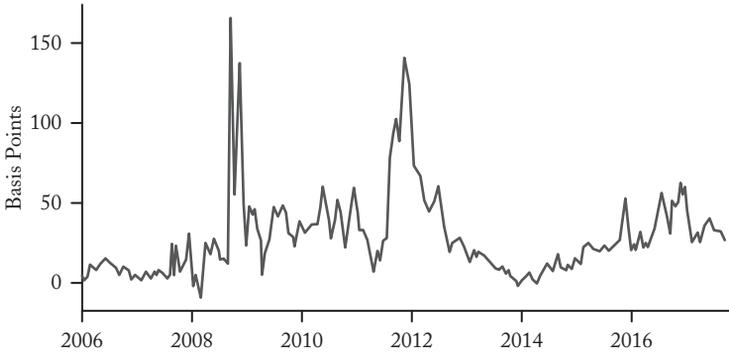
But the phenomenon of unexploited arbitrage has become more widespread and pronounced since the crisis, and derivatives markets are not always involved. We'll examine a set of anomalies in the Eurodollar, fed funds, repo, and swap markets.

One prominent example has been occurring in foreign exchange markets. The cross-currency basis is the spread between the cost of borrowing U.S. dollars directly versus indirectly, by borrowing in local currency and engaging in a foreign exchange swap. Many non-U.S. entities, such as non-U.S. banks lending dollars to their own customers, want to fund in dollars, and may even be providing dollar

¹⁹These views can be compared in a recent issue of the *Cato Journal*. The so-called neo-Fisherian view that inflation may be suppressed by low nominal rates is summarized in Bullard (2016); the view that the equilibrium real rate may be higher than a low inflation rate might lead one to believe in Borio (2016).

²⁰See Shleifer and Vishny (1997) and a more recent survey, Gromb and Vayanos (2010).

FIGURE 5
CROSS-CURRENCY BASIS OF EUR-USD, 2006–17



SOURCES: Spread between three-month USD Libor and USD borrowing rates implied by foreign exchange spot and swap markets, euro-dollar rate. Daily, smoothed using two-week moving average, in basis points. Market data from Bloomberg LP.

funding to their own customers. However, they have less access to the U.S. money markets than U.S. banks, particularly to privileged funding sources such as insured deposits, and face transaction costs in obtaining dollar funding via the foreign exchange markets, so the cross-currency basis is generally positive, but small. The basis spiked at the onset of the global financial crisis and has been unusually wide since, though it should be reduced through a low-risk near-arbitrage carried out by U.S. banks with natural access to dollar deposits. Figure 5 displays the basis for the U.S. dollar against the euro, but it has appeared for a range of major exchange rates against the dollar.²¹

Another anomaly arose almost immediately after the Fed's introduction of interest on reserves. Reserve balances at the Fed are free of credit risk, but federal funds transactions are unsecured. Banks should be unwilling to lend in the interbank market at an interest rate below that on reserves, which should therefore act, as originally intended, as a floor on the funds rate. With the money markets awash in liquidity, the fed funds rate has generally been well below IOER.

Some nondepository institutions, the housing-related government-sponsored enterprises (GSEs) Fannie Mae and Freddie Mac, and the

²¹See Du, Tepper, and Verdelhan (2017) and Borio et al. (2016).

Federal Home Loan Banks (FHLBs) may keep balances at the Fed, but may not earn interest on reserves. Currently, the limited activity in the fed funds market involves lending by the FHLBs of their large mortgage interest inflows until they are reinvested in new mortgages. The banks could borrow from the FHLBs in the funds market and receive interest on reserves at the Fed, earning a near-riskless spread.

The persistence of the spread has been attributed to regulatory changes, particularly those related to leverage, liquidity, and deposit insurance, as well as the potential stigmatization of banks seen to be borrowing heavily in the funds market. The asset-based deposit insurance fee adds costs to holding balances at the Fed for those banks subject to it. Most of the arbitrage activity is therefore carried out by FBOs, which are subject to different regulatory and accounting rules from U.S. banks, and now hold half of all reserve balances.

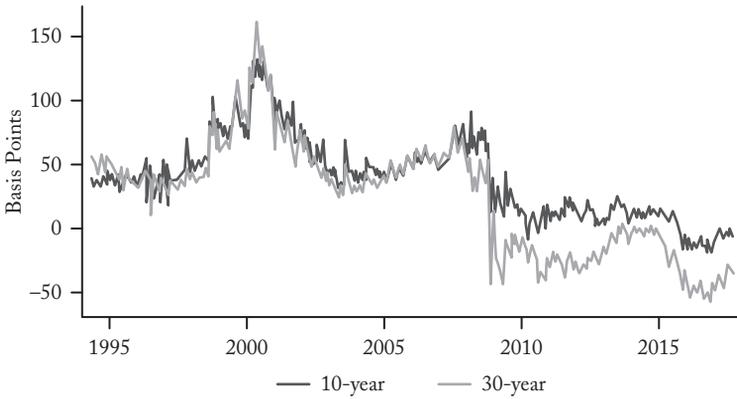
Another postcrisis anomaly is the prevalence of negative swap spreads. Interest-rate swap fixed rates are generally at least somewhat higher than on-the-run U.S. Treasury yields, since swaps have counterparty and liquidity risk premiums absent from the bond yields. Since the crisis, the 30-year spread has been negative, and more recently the 10-year spread has turned negative as well. The swap rates are driven lower in part by institutional investors' strong demand for duration risk, part of the reaching for yield phenomenon discussed earlier.²²

An arbitrage in which banks pay the fixed rate on swaps and establish government bond short positions in the repo market would bring a low-risk net positive cash flow from two sources, the spread between the swap rate and bond yield and the spread between the Libor floating rate paid on the swap and the repo financing rate of the bond short position (Figure 6).

U.S. repo markets are heavily impacted by many of the recent regulatory changes. The repo markets are also at the crux of efforts to normalize monetary policy without disruption. With the fed funds rate essentially disabled as a policy tool, the Federal Reserve has relied increasingly on repo markets to recreate scarcity in the money markets and signal and enforce its target money market rate.

²²The excess of credit default swap premiums over corporate bond spreads to plain-vanilla swap rates, described in Boyarchenko et al. (2016), is a similar phenomenon.

FIGURE 6
SWAP SPREADS, 1994–2017



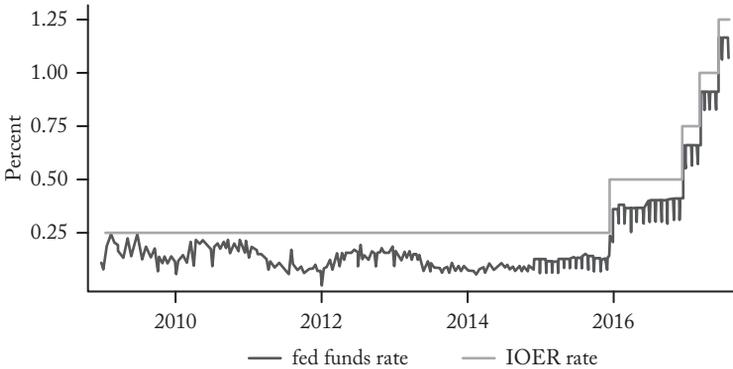
SOURCES: Spread of plain-vanilla interest-rate swaps over yield of Treasury of like maturity, basis points, daily. Bloomberg LP.

The target market-adjusted rate increasingly appears to be the GCF repo rate. The evolution dovetails with regulators' desire to replace Libor, a set of indexes prepared by banks of interbank rates of different terms to maturity for different currencies, with a more purely market-based index. GCF repo rates have been suggested as replacements for both the fed funds rate as a policy rate and for Libor as a loan and derivative benchmark.²³

Repo markets have displayed several anomalies in recent years. Repo lending is secured, and should therefore be executed at lower rates than funds trades. Repo rates have instead generally been above the funds rate in recent years, although the spread has diminished since the end of 2016 (Figure 7). Repo rates have also been quite volatile. Accounting and regulatory reporting constraints have long induced month- and quarter-end "window dressing" spikes in money market rates. These have grown more pronounced for GCF repo. European banks are subject to quarter-end rather than daily average

²³See Committee on the Global Financial System (2015) for an overview of anomalies in money markets in relation to the implementation of monetary policy in several advanced economies. Federal Reserve Bank of New York (2017) provides an update for the U.S. as part of its discussion of the effectiveness with which policy rate increases have been passed through to money markets generally.

FIGURE 7
FEDERAL FUNDS AND INTEREST ON
RESERVES RATE, 2008–17



SOURCES: Federal funds effective rate and interest rate on reserve balances at Federal Reserve Banks. Daily, percent, from December 18, 2008. Bloomberg LP.

reporting, and are particularly inclined to withdraw from lending on reporting dates. Quantitative restrictions, in which intermediaries show reluctance to accept investors' cash near reporting dates, are reportedly also appearing.

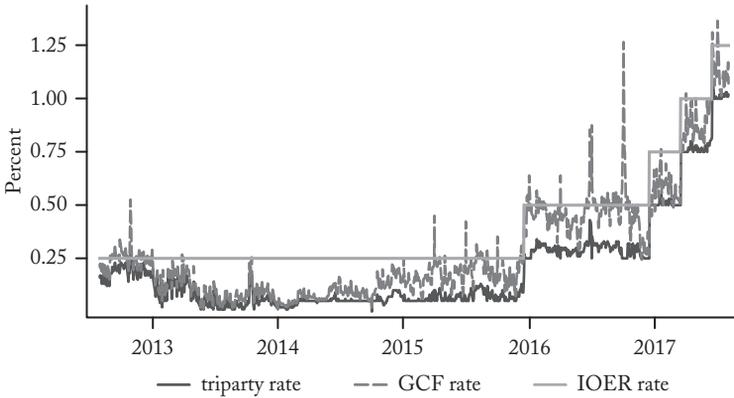
GCF repo trades at a persistent spread above triparty repo. Triparty repo can be viewed as the marketplace in which the dealers seek funding for this business from MMMFs and other providers of short-term wholesale funding, while GCF repo can be viewed as the marketplace in which dealers trade among themselves and offer repo to hedge funds and other market participants taking positions in securities. In this sense, the GCF-triparty repo spread has been likened to the bid-ask spread of the repo market. This spread is wide and fluctuates quite a bit, particularly at month-end. It ought to be contained by arbitrage, since the trades through which the spread is earned are low in risk (Figure 8).²⁴

Since the crisis, the incidence of fails in the repo market has greatly increased (Figure 9).²⁵ Fails occur when borrowers of securities don't

²⁴Greenwood, Hanson, and Stein (2016) attribute negative swap and the wide GCF-triparty spread to the SLR.

²⁵The figure omits the extremely large increase in fails during the worst part of the crisis in 2008 and 2009.

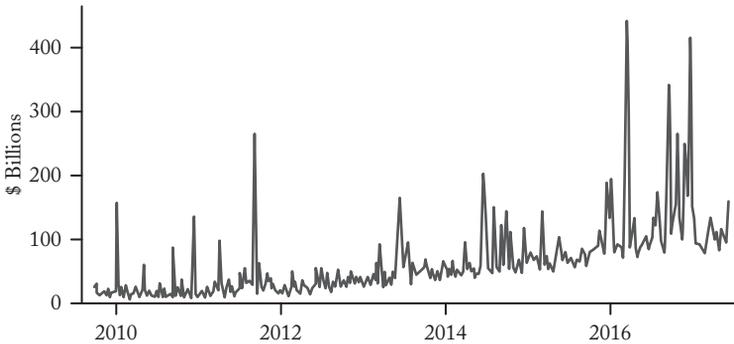
FIGURE 8
TRIPARTY AND GCF REPO RATES, 2012–17



SOURCES: Triparty: BNY Mellon Treasury Tri-Party Repo Index (Bank of New York Mellon). GCF: DTCC GCF Repo Index for Treasury (Bloomberg LP).

return them to the lenders and take back the cash they lent, therefore forgoing interest on it. They can occur for operational reasons, or because there are large short positions in the specific securities borrowed, among other reasons, and are a longstanding phenomenon in repo markets. The incentive to fail to deliver is greater with low

FIGURE 9
TREASURY FAILS, 2009–17



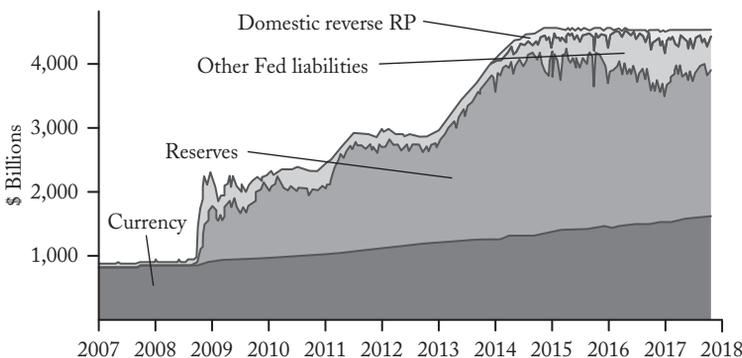
SOURCES: Total of fails to receive and to deliver, weekly, from September 30, 2009, \$ bill. Federal Reserve Bank of New York, Primary Dealer Statistics.

interest rates, since the interest forgone on the cash lent is small. A fails penalty was introduced in 2010, but fails have increased greatly nonetheless. Occasionally, they spike briefly to extreme levels. The increase may be due not only to low interest rates but also to the large share of Treasury securities now held on the Fed's balance sheet and thus out of the market.

Some of the anomalies we've looked at have become somewhat attenuated during 2017, though they haven't disappeared or returned to the precrisis state. One possible reason may be a decline in the Treasury's General Account deposit balances as the federal debt limit nears. Since the asset side of the balance sheet is being held constant as a matter of policy until the pace of reinvestment is slowed as part of normalization, another Fed liability position must rise. Banks' reserve balances, which had been declining since 2014, have increased moderately, adding to the supply of liquidity (Figure 10).

Arbitraging these anomalies requires considerable balance sheet and in the presence of a potentially binding leverage ratio, it has been argued, additional balance sheet has a high cost of capital. But the arbitrage is quite low in risk, so the hurdle rates should not be assumed frozen at high levels. Adding very low-risk activities should

FIGURE 10
FEDERAL RESERVE LIABILITIES, 2007–17



NOTES: Currency includes other liabilities and capital. Other Fed liabilities include foreign official reverse repos, deposits other than reserves, and Treasury cash.

SOURCE: Federal Reserve Board, H.4.1 release, Table 1.

lower them.²⁶ Rates of return bank equity investors currently seek are based on current equity levels and asset mix. Just as relying more on equity funding would induce a lower market-clearing return on equity, so also would a lower-risk asset mix. The argument that the regulatory capital and liquidity constraints raise the required return on capital above the hurdle rate doesn't therefore fully explain the phenomenon. One explanation may be debt overhang, a situation in which firms are so highly levered that at least part of the return to profitable equity-financed activities accrues to debt holders.

Conclusion

"Inadequate liquidity" may not be the best way to describe the phenomena we currently see. The anomalies and conflicts we've described indicate a larger underlying disfunction. Each anomaly is hard to trace back to a specific regulatory change. Not only have there been major regulatory changes, but the monetary policy response to the crisis, the crisis itself, the continuation of precrisis trends, and the financial industry's adaptation to these also influence market functioning. But taken as a whole, they indicate a general decline in market responsiveness.

Liquidity currently seems ample, but perhaps only because market participants don't urgently need it right now. The market appears persistently less able to withstand large shocks. Disruptive shocks are likely in the years to come, especially sudden changes in expectations that increase the desire for liquidity, but also cause it to disappear. There are many potential sources that have been of concern for some time: policy normalization, conflict risks, and problems in a number of specific large countries. Overseas reliance on U.S. dollar funding is as great as ever, while the intermediation channels by which it is obtained remain as fragile as ever.

The slowness and difficulty of arbitrage is important. While slow arbitrage is normal, and exceptionally slow arbitrage was part of the crisis, arbitrage is a key mechanism through which shocks and surprises are absorbed. If it's impaired, the ability to adjust is impaired, and disruptions are more likely to take the form of large price swings, reductions in credit, and other typical crisis phenomena. The banking

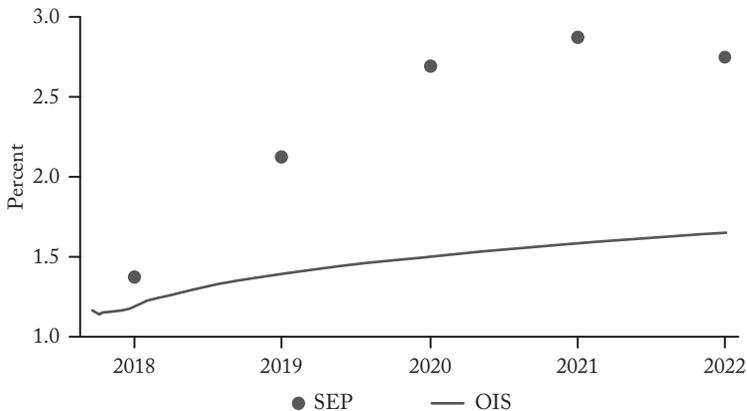
²⁶The argument is analogous to that in favor of much higher equity ratios (see Admati et al. 2013).

system remains at the core of the arbitrage process, but is still highly leveraged and dependent on public-sector support in severe stress.

One potential source of shocks and surprises is the anticipated normalization of monetary policy and gradual reduction in central bank balance sheets or in the pace of new purchases. There are potential disruptions if normalization goes forward as expected, and others that might result from much faster or slower normalization than expected. Even in the absence of major surprises in the pace of normalization, the Fed's new tools for raising rates might not work as expected, resulting for example in money market rates failing to rise with higher targets, or opening wider spreads among money markets. This has not been the case for the first 75 basis points of hiking; market rates have risen in tandem with policy rates, and the ON RRP floor has held, albeit with larger volumes than initially anticipated. But it remains a risk.

The contrast between market expectations of future rates and the expectations reported by Federal Reserve officials parallels the contrast between indicators of risk aversion and risk appetite. Figure 11

FIGURE 11
FEDERAL RESERVE AND MARKET VIEWS OF FUTURE
INTEREST RATES



SOURCES: Dots represent median of Federal Open Market Committee participants' projections of future fed funds rate from the September 20, 2017, Summary of Economic Projections (SEP). Solid plot represents rates on overnight interest swap (OIS) settling on different future dates, as of September 20, 2017. Horizontal axis displays horizon of SEP rate projections and OIS settlement dates. Both from Bloomberg LP.

displays future overnight rates from a recent Summary of Economic Projections (SEP), alongside an implied forecast of the overnight rate drawn from money market derivatives markets. Several years out, the gap between market expectations and Fed projections is well over 200 basis points. The projected pace of increases and longer-term market rates have both fallen since the SEP was introduced, but the size of the gap has been remarkably persistent.

This tension will have to be resolved somehow, and it may not be resolved without market disruption. Normalization is needed, but itself poses risks. Future rates closer to Fed projections would surprise markets, while future rates closer to those implied in markets would validate their pessimism. Surprises in either direction could result from the use of new operational tools in drastically transformed markets. Low interest rates may also move discontinuously, amplifying the extensive revaluations that can be expected to accompany a rise in rates.

Low interest rates and volatility, sustained over a long period time, are apt to contribute to a further tightening of financial conditions in response to adverse shocks. The paradox of volatility suggests that hard-to-detect imbalances are growing that are apt to correct themselves disruptively, perhaps as pockets of leveraged investing turn sour. Microeconomic anomalies and evidence of market fragility increase the likelihood of disruptive market events with surprising loci and timing, exacerbating the effects of rising rates and volatility.

References

- Aaron, M.; Demers, A.; and Durr, S. (2015) “The Effect of Regulatory Changes on Monetary Policy Implementation Frameworks.” *Bank of Canada Review* (Autumn): 39–49.
- Acharya, V. V., and Schnabl, P. (2010) “Do Global Banks Spread Global Imbalances? Asset-Backed Commercial Paper during the Financial Crisis of 2007–09.” *IMF Economic Review* 58 (1): 37–73.
- Admati, A. R.; DeMarzo, P. M.; Hellwig, M. F.; and Pfleiderer, P. (2013) “Fallacies, Irrelevant Facts, and Myths in the Discussion of Capital Regulation: Why Bank Equity Is Not Socially Expensive.” Working Paper No. 2065 (October), Stanford Graduate School of Business.
- Adrian, T.; Fleming, M.; Shachar, O.; and Vogt, E. (2017) “Market Liquidity after the Financial Crisis.” Staff Report No. 796 (June), Federal Reserve Bank of New York.

- Baklanova, V.; Copeland, A.; and McCaughrin, R. (2015) “Reference Guide to U.S. Repo and Securities Lending Markets.” Staff Report No. 740 (December), Federal Reserve Bank of New York.
- Bao, J.; O’Hara, M.; and Zhou, A. (2016) “The Volcker Rule and Market-Making in Times of Stress.” Finance and Economics Discussion Series No. 2016–102, Board of Governors of the Federal Reserve System.
- Basel Committee on Banking Supervision (2011) *Basel III: A Global Regulatory Framework for More Resilient Banks and Banking Systems* (June), Bank for International Settlements.
- _____ (2013) *Basel III: The Liquidity Coverage Ratio and Liquidity Risk Monitoring Tools* (January), Bank for International Settlements.
- Bollerslev, T.; Tauchen, G.; and Zhou, H. (2009) “Expected Stock Returns and Variance Risk Premia.” *Review of Financial Studies* 22 (11): 4463–92.
- Borio, C. (2016) “Revisiting Three Intellectual Pillars of Monetary Policy.” *Cato Journal* 36 (2): 213–38.
- Borio, C.; McCauley, R.; McGuire P.; and Sushko, V. (2016) “Covered Interest Parity Lost: Understanding the Cross-Currency Basis.” *BIS Quarterly Review* (September): 29–41.
- Boyarchenko, N.; Gupta, P.; Steele, N.; and Yen, J. (2016) “Trends in Credit Market Arbitrage.” Staff Report No. 784 (July), Federal Reserve Bank of New York.
- Bullard, J. (2016) “Permazero.” *Cato Journal* 36 (2): 415–29.
- Caballero, R. J.; Farhi, E.; and Gourinchas, P. (2017) “The Safe Assets Shortage Conundrum.” *Journal of Economic Perspectives* 31 (4): 29–46.
- Calomiris, C. W. (2009) “A Recipe for Ratings Reform.” *The Economists’ Voice* 6 (11): 1–4.
- Committee on the Global Financial System (2015) “Regulatory Change and Monetary Policy.” CGFS Paper No. 54 (May), Bank for International Settlements.
- European Commission (2017) *Drivers of Corporate Bond Market Liquidity in the European Union*. Available at https://ec.europa.eu/info/sites/info/files/171120-corporate-bonds-study_en.pdf.
- Federal Reserve Bank of New York (2017) *Domestic Open Market Operations during 2016*. Available at www.newyorkfed.org/medialibrary/media/markets/omo/omo2016-pdf.

- Fleckenstein, M.; Longstaff, F. A.; and Lustig, H. (2014) “The TIPS-Treasury Bond Puzzle.” *Journal of Finance* 69 (5): 2151–97.
- Foucault, T.; Pagano, M.; and Roell, A. (2013) *Market Liquidity: Theory, Evidence, and Policy*. New York: Oxford University Press.
- Greenwood, R.; Hanson, S. G.; and Stein, J. C. (2015) “A Comparative-Advantage Approach to Government Debt Maturity.” *Journal of Finance* 70 (4): 1683–1722.
- _____ (2016) “The Federal Reserve’s Balance Sheet as a Financial-Stability Tool.” Economic Policy Symposium Proceedings, Federal Reserve Bank of Kansas City: 335–97.
- Gromb, D., and Vayanos, D. (2010) “Limits of Arbitrage.” *Annual Review of Financial Economics* 2: 251–75.
- Ihrig, J. E.; Meade, E. E.; and Weinbach, G. C. (2015) “Rewriting Monetary Policy 101: What’s the Fed’s Preferred Post-Crisis Approach to Raising Interest Rates?” *Journal of Economic Perspectives* 29 (4): 177–98.
- International Organization of Securities Commissions (2017) “Examination of Liquidity of The Secondary Corporate Bond Markets.” Available at www.iosco.org/library/pubdocs/pdf/IOSCOPD537.pdf.
- Kim, S.; Plosser, M. C.; and Santos, J. A. C. (2017) “Macprudential Policy and the Revolving Door of Risk: Lessons from Leveraged Lending Guidance.” Staff Report No. 815 (May), Federal Reserve Bank of New York.
- Mizrach, B. (2015) “Analysis of Corporate Bond Liquidity.” Research Note, FINRA Office of the Chief Economist. Available at www.finra.org/sites/default/files/OCE_researchnote_liquidity_2015_12.pdf.
- New England Asset Management (2017) “2016 Investment Highlights: Historic and Future.” *Perspectives* (July).
- Shleifer, A., and Vishny, R. W. (1997) “The Limits of Arbitrage.” *Journal of Finance* 52 (1): 737–83.
- _____ (2011) “Fire Sales in Finance and Macroeconomics.” *Journal of Economic Perspectives* 25 (1): 29–48.

REORGANIZATION OF FAILING FINANCIAL FIRMS: A CAPITAL STRUCTURE SOLUTION

Steven Gjerstad

The underlying rationale of joint-stock companies is that equity holders bear the largest proportional risk of an enterprise, and are rewarded with the most control over the firm and all of its upside risk. Equity owners are lowest in the debt hierarchy: their claims come after all other legitimate claimants of a defaulted enterprise. However, they are the highest in the hierarchy of control: in principle they select the directors of the enterprise who in turn select its management. This article argues that the rationale behind equity owner control should be extended in the debt and control hierarchies at least one more level. The unsecured debt of a firm should have a junior-most tranche, and control of the firm should devolve on the owners of that debt automatically and immediately upon the failure of the firm. Although the reorganization proposal described in this article would function effectively for any failing enterprise, this article considers the important and challenging problem of restructuring a failing systemically important financial institution (SIFI).

Legal Aspects of Existing and Proposed Legislation

Currently, there are two approaches to the reorganization of a SIFI in the United States—both mandated under the Dodd-Frank

Cato Journal, Vol. 38, No. 1 (Winter 2018). Copyright © Cato Institute. All rights reserved.

Steven Gjerstad is a Presidential Fellow at Chapman University's Economic Science Institute.

Wall Street Reform and Consumer Protection Act—and a proposed modification to these approaches in the Financial CHOICE Act of 2017, passed by the House of Representative in June 2017.¹

Title I of the Dodd-Frank Act requires each systemically important bank to develop detailed plans—“living wills”—that specify procedures for its own resolution if it is about to collapse. The resolution strategies submitted to the Federal Reserve and the Federal Deposit Insurance Corporation (FDIC) to date rely primarily on either (1) a holding company that will spin off its subsidiaries while the holding company itself files for bankruptcy protection or (2) a “bridge bank” that will receive the high-quality assets of the failing firm and many of its liabilities, leaving most impaired assets and the remainder of its liabilities behind with the parent firm that enters bankruptcy. These two approaches are known jointly as the single-point-of-entry (SPOE) approach.²

The more complex banks, with numerous subsidiaries, will be structured with a bank holding company that would enter bankruptcy in the event of a failure. The resolution plans specify that subsidiaries should be well-capitalized, with sufficient liquid assets and with strong balance sheets, prior to a bankruptcy filing by the parent holding company. During bankruptcy proceedings the bank’s major subsidiaries would continue operations until they are sold off and the proceeds of the sales are returned to the estate of the bankrupt holding company. The second approach would be to form a bridge bank that would receive most of the high-quality assets of the failing bank and a portion of its liabilities. The impaired assets of a failing bank and the remainder of its liabilities would be left behind in the bankrupt firm. Both of these approaches are at odds with a fundamental precept of

¹Throughout this article, the Dodd-Frank Wall Street Reform and Consumer Protection Act will be referred to as the “Dodd-Frank Act” (see www.gpo.gov/fdsys/pkg/BILLS-111hr4173enr/pdf/BILLS-111hr4173enr.pdf). The Financial CHOICE Act of 2017 will be referred to as the “CHOICE Act.” The parts of the CHOICE Act that repeal and replace Title II of the Dodd-Frank Act are Title I, Subtitle A (repeal) and Title I, Subtitle B (replace). See https://financialservices.house.gov/uploadedfiles/hr_10_the_financial_choice_act.pdf.

²See Title I, Section 165 (d) of the Dodd-Frank Act and the public submissions by banks and the responses from the Federal Reserve and the FDIC at www.federalreserve.gov/bankinforeg/resolution-plans.htm. See Lee (2015) for analysis of the banks’ resolution plans, especially pp. 464–86, which describe the adoption of the holding company and bridge bank approaches.

American bankruptcy law: similarly situated creditors should be treated similarly in the settlement of the bankrupt firm's affairs.

Under the holding company structure, the creditors of the subsidiaries that are sold off will have claims on going concerns with positive net values, so their claims will be unimpaired. On the other hand, the holding company will be bankrupt, and it will absorb all of the losses of the institution. Similarly, creditors of a bridge bank would have claims against a going concern, while the creditors whose claims are assigned to the parent firm will have claims against a bankrupt firm. The creditors of the bankrupt parent company, unlike the creditors of subsidiaries or a bridge bank, will be required to absorb all of the losses. Unequal treatment of similarly situated creditors is likely. The only way to avoid it would be to place all of the unsecured debt of the firm in the bankruptcy estate. Nothing in the Dodd-Frank Act requires or suggests that this will be done, and nothing would prevent the firm from distributing its unsecured debt to subsidiaries or a bridge company according to its own preferences. This opens the door to both unequal treatment of similarly situated creditors and to favoritism toward some creditors by the management of the firm in the lead-up to its collapse.

If the resolution process devised by the company in compliance with Title I, Section 165(d) of the Dodd-Frank Act is not implemented successfully and in a timely manner, Title II requires the FDIC to take control of the firm in receivership.³ Title II vests almost unlimited authority in a receiver appointed by the FDIC.⁴ Criticism has been leveled at the Dodd-Frank Act for its harshness⁵

³Even the FDIC has sought to avoid the path of liquidation, with most of its efforts directed toward review and assessment of the banks' Title I resolution plans. See Lee (2015: 476–78) for discussion of the pivot by the FDIC from Title II liquidation toward Title I resolution planning.

⁴McDermott and Turetsky (2011: 412) describe the extent of the receiver's control. "Once the FDIC is appointed receiver of a covered financial company, it assumes virtually complete control over the company and the receivership process. The perfunctory role of the courts in the core receivership process ends, and there are limited avenues for challenging the various ancillary decisions that the FDIC may make in pursuing the liquidation."

⁵McDermott and Turetsky (2011: 404) argue that "the provisions of the Act and the powers delegated to the FDIC and other government authorities may be draconian when implemented. The right to decide whether to initiate receivership proceedings is vested in government authorities, not in financial companies' boards, management, or stakeholders, and is subject only to very limited judicial review that is highly deferential to such authorities."

and for its subversion of constitutionally guaranteed protections.⁶ Other legal arguments could be mounted against Dodd-Frank, such as those developed in Hamburger (2014) and in Lawson (2015), to address the expansive authority of administrative agencies. Hamburger (2014) argues that agencies frequently combine legislative functions in their rulemaking, executive functions in their oversight and enforcement, and judicial functions with their administrative law judges. One of Hamburger's primary arguments is that agencies violate the separation of powers, remove legislative functions from elective bodies, and eliminate review by an independent judiciary. For example, Hamburger (2014: 257) argues that

administrative procedure . . . is justified on the ground that the courts are the real target of constitutional guarantees of due process and other procedural rights. It is not plausible, however, to suggest that administrative tribunals are less confined by procedural guarantees than the courts, or that administrative process satisfies the due process of law. Procedural rights developed . . . precisely to bar extralegal adjudication. And the growth of administrative adjudication only confirms the importance of procedural rights as limits on extralegal power. Rather than satisfy the due process of law, administrative process is exactly what the guarantee of due process forbids.

These arguments clearly have relevance to agency authority in Dodd-Frank, which is extensive and operates explicitly without even the administrative law review provided by other agencies, such as the SEC, the FDIC, the NLRB, the FAA, the EEOC, and dozens of other federal agencies. Recent cases, such as *Bandimere v. SEC* in the 10th Circuit Court of Appeals and *Burgess v. FDIC* in the 5th Circuit Court of Appeals, indicate a growing discomfort within the judiciary toward quasi-judicial review within an administrative agency of that agency's own proceedings. Although the *Bandimere* and *Burgess* decisions are confined to the question of how administrative law judges are appointed, it would hardly seem to solve the problem of the irregularity of those proceedings by eliminating even agency review.

⁶Merrill and Merrill (2014) argue that Dodd-Frank Title II violates the due process clause in the Fifth Amendment and subverts judicial review in Article III of the Constitution, among other issues.

The legal arguments against Dodd-Frank are worth pursuing in their own right, but the argument put forth and emphasized in this article is that the procedures in Dodd-Frank's Title II impose unnecessary administrative oversight and authority in the resolution of distressed financial firms: a simpler, less disruptive, more transparent market-oriented process can be implemented that defines and respects rights of the firms' creditors.

The challenge of a reorganization regime for SIFIs is to balance the need to maintain their core functions without interfering with the goal of giving equal treatment to similarly situated creditors. Dodd-Frank's Title II fails explicitly in that regard. For example, in Title II, Section 210(h)(1)(B): "Upon the creation of a bridge financial company . . . such bridge financial company may . . . assume such liabilities . . . as the Corporation may, in its discretion, determine to be appropriate." This places the FDIC squarely in the role of partitioning creditors into one group that will experience no loss and another group that will bear all of the losses of the bankrupt firm.

Title I of the Dodd-Frank Act is no better: the management of the failing firm will determine the partition. The possibility of favoritism or even coercion and retribution is apparent. Creditors of a firm that demand payments due from a faltering bank could be asked to back off or even lend more to the failing bank. If the creditors refuse, the management of the failing firm could put all of the debts of that creditor into the bankruptcy estate rather than the bridge bank. The possibilities for manipulation open to a failing bank are limited only by the imagination of the management and board of directors. One of the purposes of bankruptcy is to explicitly close avenues for fraudulent conveyance. By avoiding bankruptcy proceedings, Title I opens endless avenues for mischief by desperate or even vengeful management of failing firms.

The CHOICE Act places the reorganization process back in the hands of bankruptcy courts, but it falls back on the same device of a bridge bank that continues as a going concern, and a parent bank that enters bankruptcy. The apparent motive of this legislation is to regularize the legal framework for SIFI reorganization. While the CHOICE Act restores the role of the courts, as with both Title I and Title II reorganization procedures, some creditors would be favored by becoming creditors of the bridge bank, and others would be in the unfortunate position of being creditors of the bankruptcy estate.

The proposal offered in this article, based on research conducted with Vernon Smith, would maintain the core functions of the firm through reorganization yet avoid the different treatment of similarly situated creditors by creating a class of creditors who know *ex ante* that they would be called upon to absorb losses.⁷ The challenge of the reorganization of a SIFI is that it must take place quickly—and for the stability of the financial system, core functions of the failing firms should be placed in a solvent entity that operates without interruption. That is the rationale for creating a bridge bank or assuring that major operations continue within key subsidiaries of the parent firm. That necessity implies that some creditors will absorb losses. Our proposal simultaneously assures the solvency of the firm by assigning losses to one predesignated tranche of creditors, but it compensates those creditors with ownership of the firm after reorganization. Notably, under both Dodd-Frank reorganization approaches, as well as under the CHOICE Act proposal, the creditors who absorb the losses are not compensated in any way for their loss, while control devolves onto parties selected by the management of the failing firm under Title I and it devolves onto the FDIC-appointed receivers under Title II.

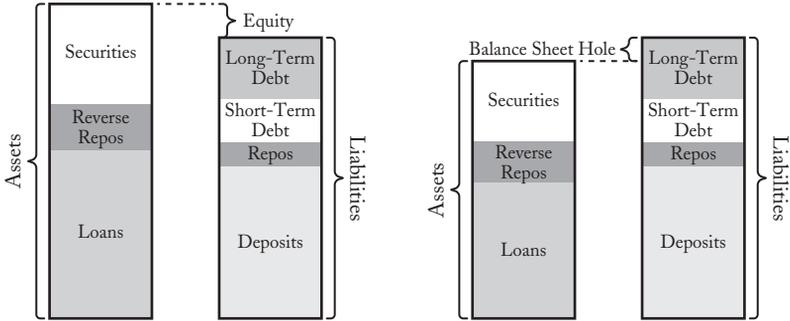
After a brief review of the balance sheet consequences of an asset collapse, I provide a summary of reorganization bonds, which are a market-oriented alternative resolution process that respects the property rights of creditors, preserves as much of the value of a failing firm as possible, and minimizes disruption to credit markets, asset markets, and the economy.

Asset Collapse and Insolvency

The left side of Figure 1 shows a stylized balance sheet with key asset and liability items for a bank with positive equity. The right side shows the same bank's balance sheet if the value of its assets collapses—that is, the bank's equity position disappears and the bank becomes insolvent. A bailout would add cash to the asset side of the balance sheet and more debt to the liability side of the balance sheet, leaving the balance sheet hole the same size. A bailed-out bank typically takes earnings from the healthy portion of its asset portfolio as they come in over the course of many years and moves them to loss

⁷For a summary of our proposal, see Gjerstad and Smith (2014).

FIGURE 1
 ASSETS AND LIABILITIES BEFORE (LEFT) AND AFTER
 (RIGHT) ASSET VALUES COLLAPSE



reserves so that impaired assets can be removed from the books. During this long period the bank has suppressed earnings to dedicate to dividends, stock buybacks, or organic growth. One consequence of this is that the bank has difficulty finding capital investment, and it will typically deleverage as the only avenue available to it to raise its capital to asset ratio. A method is needed to remove some of the liabilities from the bank’s balance sheet. Bankruptcy does that, but at the cost of major disruption to the firm, to financial markets, and to the wider economy. Our proposal achieves the same goal with far less trauma.

Reorganization Bonds

Gjerstad and Smith (2014) propose the creation of a class of bonds that sits between equity and all other creditors in the hierarchy of firm obligations. These bonds, which we called “reorganization bonds” (or “R bonds”), would be converted to equity immediately upon failure of a firm.⁸ I want to emphasize that R bonds differ in a crucial respect from contingent convertible bonds, because we propose that ownership and control of the corporation would pass to the

⁸By a failure, I mean events like those that precipitated government interventions on behalf of Bear Stearns between March 14 and March 16, 2008, the collapse of Lehman Brothers six months later, and the intervention on behalf of AIG on September 16, 2008.

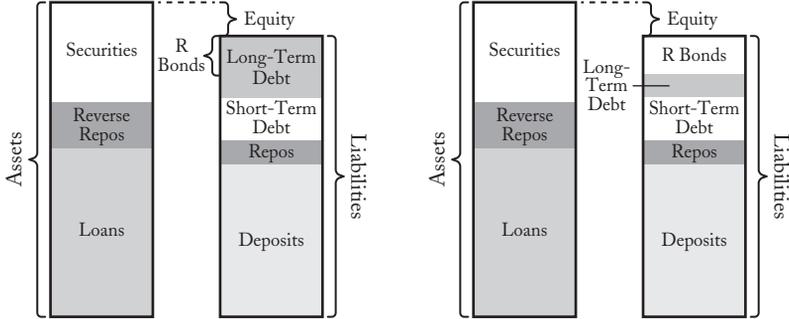
bondholders with conversion of the bonds to equity. This is crucial because contingent convertible bonds simply provide the managers who have failed with a new pool of capital, and profits that the firm accrues after conversion would be shared by the owners of the contingent convertible bonds with incumbent shareholders.⁹ Under our proposal, if an asset value collapse causes a firm to run out of equity capital—that is, the firm becomes insolvent—the incumbent equity holders' shares are eliminated, the R bonds are converted to equity, and the holders of the equity that was created by conversion of the R bonds become the sole owners of the firm.

Figure 2 shows how R bonds are created from standard long-term debt. Figure 3 shows the effect of eliminating the claims of the original equity holders and converting the R bonds to a new equity pool. After the R bonds are converted, the board of directors selected by the bondholders' committee and the new management team will take control of the bank. Contingent convertible bonds would reward mismanagement and failure with new capital; our proposal circumvents that incentive flaw.

With this structure, R bonds should trade at almost the same price as the bank's standard corporate bonds when the likelihood of failure is near zero, so that a sound bank will face no additional cost of funding with this arrangement. If the bank does become insolvent, then prior equity holder claims are eliminated when R bonds convert to equity. If losses on assets are less than the sum of the precrisis book value of equity and the amount of R bonds, then the bank's solvency is restored under our procedure and the reorganization will produce

⁹Many authors suggest that these bonds could convert to equity if the market capitalization of the firms passes below some threshold, such as 2 percent of the firm's liabilities. When the threshold is crossed, some amount of the bonds would be converted to equity at the market price of equity shares when the threshold was passed. There are at least three significant problems with this approach. One issue is that if the equity price crosses the threshold on its way lower, bond conversion bails out the incumbent shareholders before their shares lose more or all of their value. The second consideration is that these procedures also leave the incumbent management and board in control of the firm. A third problem is that, with the management and board in control of the firm, the new equity holders have minimal protection against the incumbent management and board enriching themselves at the expense of the new equity holders before the new equity holders can wage a fight for control of the firm. For all of these reasons I argue that it would be better to allow the firm to enter a prepackaged bankruptcy where the bondholders are rewarded with control of the firm.

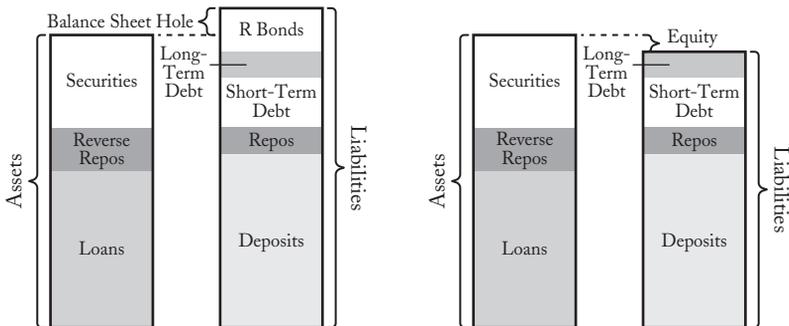
FIGURE 2
SPLITTING LONG-TERM DEBT INTO R BONDS AND
STANDARD LONG-TERM BONDS



a new equity cushion owned exclusively by the investors whose R bonds were converted. With established criteria that trigger conversion, the ad hoc nature of bailouts, the public funds that support them, and their severe incentive distortions can all be eliminated.

Conversion of long-term bonds to equity is a key element of our proposal. We propose that, if equity capital is depleted and the bank is on the brink of default, 12 percent of the bank's liabilities could be converted from debt into equity capital. Some of the R bond debt so converted will absorb the balance sheet hole that preceded the bank's collapse, but if that hole is less than 12 percent of the bank's

FIGURE 3
CONVERTING R BONDS TO EQUITY AFTER AN
ASSET COLLAPSE



liabilities, then the bank will have a positive equity capital position after conversion.

Table 1 shows the liabilities of the ten largest banks in the United States at the end of 2007; it also shows the amount of long-term debt that each one had outstanding. Each of these firms had between 12.6 percent and 26.4 percent of its liabilities in the form of long-term bonds, so that each would have been able to meet the threshold for R bond issuance without a substantial change to its financing costs.

At the end of 2007, the book value of equity capital at Bear Stearns, Lehman Brothers, Merrill Lynch, and Morgan Stanley had fallen to levels just above 3 percent of liabilities. (See Table 1, columns 4 and 5.) The Federal Reserve carried out a back-door bailout of Bear Stearns by lending \$28.82 billion to Maiden Lane LLC to purchase impaired Bear Stearns assets in March 2008. Lehman Brothers failed in September 2008. Merrill Lynch was sold to Bank of America with loss guarantees on \$118 billion of its

TABLE 1
LONG-TERM DEBT AND EQUITY RELATIVE TO TOTAL
LIABILITIES AT THE END OF 2007

	Total Liabilities	Long-Term Debt	Long-Term Debt to Liabilities	Book Value of Equity	Equity to Liabilities
Citigroup	\$2,074.03	\$427.11	20.6%	\$113.60	5.5%
Bank of America	\$1,568.94	\$197.51	12.6%	\$146.80	9.4%
JPMorgan Chase	\$1,438.93	\$199.01	13.8%	\$123.22	8.6%
Goldman Sachs	\$1,076.97	\$164.17	15.2%	\$42.80	4.0%
Morgan Stanley	\$1,014.14	\$190.62	18.8%	\$31.27	3.1%
Merrill Lynch	\$988.12	\$260.97	26.4%	\$31.93	3.2%
Wachovia	\$702.69	\$161.01	22.9%	\$76.87	10.9%
Lehman Brothers	\$668.57	\$123.15	18.4%	\$22.49	3.4%
Wells Fargo	\$527.81	\$99.39	18.8%	\$47.63	9.0%
Bear Stearns	\$383.57	\$68.54	17.9%	\$11.79	3.1%
Total	\$10,443.77	\$1,891.48	18.1%	\$648.41	6.2%

NOTE: Dollar figures in billions.

SOURCE: Total liabilities, long-term debt, and book value of equity from firms' Annual Reports.

assets from the Federal Reserve. All three of these firms had inadequate book value of equity capital by the end of 2007, and it is quite likely that the asset sides of their balance sheets were overstated due to overvalued mortgage securities on their books.¹⁰ By January 8, 2008, the market capitalization of Bear Stearns was only \$8.4 billion, or about 2.2 percent of its liabilities. This is a threshold that clearly requires a contingency plan for reorganization. By July 14, 2008, Lehman Brothers' market capitalization had fallen as low as \$8.6 billion. With \$613 billion in liabilities, equity capital had fallen to 1.4 percent of liabilities. On the same day, market capitalization of Washington Mutual fell to \$5.5 billion, or 1.9 percent of its liabilities. With these firms in such dire condition, under our proposal bondholders' committees would begin final preparations for a takeover.

Challenges for Resolution Plans and Orderly Liquidation

The Title I Resolution plans submitted to the Federal Reserve and the FDIC have a number of common features. Some of the banks have implemented a holding company structure. Others would utilize a bridge bank in the event of failure. When a holding company is utilized, it would commit substantial resources to its subsidiaries before it enters bankruptcy. The subsidiaries would continue as going concerns until they are sold and the proceeds are returned to the estate of the holding company. A bridge bank would operate similarly, but may be preferred for a firm that has few subsidiaries.

Despite the limitations of the Title I resolution procedures, assessments of the banks' plans by the FDIC and the Federal Reserve Board have led to progress on several important problems. Resolution planning has required firms to better align their business operations with their material legal entities, insure that the constituent parts of the firm have access to financial market utilities (especially payment systems and trading and settlement operations), and the banks have implemented service agreements that would

¹⁰Washington Mutual provides a good example of this. In its 10-Q filing on June 30, 2008, Washington Mutual had book value of equity capital of \$26.09 billion (about 9.2 percent of its liabilities), but at that time the market value of its equity capital was only \$8.24 billion (about 2.9 percent of equity capital). A closer examination of its balance sheet reveals an extraordinary level of exposure to residential real estate, with \$230.2 billion in real estate loans and \$19.2 billion in mortgage-backed securities. Real estate lending amounted to 79 percent of assets.

provide continued access to information technology support and other critical services to subsidiaries upon a breakup of the firm. Significant progress has also been made to extend the automatic stay in bankruptcies on contracts to include derivatives, futures, and swaps. This makes less likely a repeat of the severe losses precipitated by collateral seizures and sales that followed the Lehman bankruptcy. But the banks' resolution plans depend almost entirely on the optimistic assumption that Global Systemically Important Financial Institutions (G-SIFIs) can resolve themselves and avoid disruption to the U. S. and global financial system by breaking themselves apart and selling the pieces—presumably to other G-SIFIs. Of course, financial distress is highly correlated, so many of the largest financial firms will be turning to one another to raise capital. Even if that were possible, the sale of Lehman Brothers investment banking and capital market accounts to Barclays in a Section 363 sale in September 2008 provides insight into the limited capital raised by such sales, and the potential for serious harm to the interests of the seller—in this case the Lehman Brothers estate. According to the Trustee of the estate, James Giddens, Lehman Brothers transferred assets worth approximately \$11,869 million to Barclays. In addition, Barclays gained over 72,000 customer accounts with assets of \$43,000 million (about \$600,000 per account).¹¹ Barclays in return paid \$2,438 million on behalf of Lehman Brothers to third parties.

Title II of the Dodd-Frank Act would transfer control over a faltering financial institution to the FDIC as the receiver, charged with liquidation of the firms' assets and payment of claims against the firm. The draconian elements of Title II are well-known,¹² but Title II is also set up to wreak economic havoc, since it relies solely on sales of major business lines and liquidation of the firm. The Lehman

¹¹See the "State of the Estate," pp. 24–25, August 16, 2016, from Hughes Hubbard & Reed (the law firm of the trustee).

¹²McDermott (2010: 2), in an analysis from Skadden, Arps, Slate, Meagher & Flom LLP argues that "the potential harshness of the Act ultimately may mean that its most salutary effect will be to minimize the circumstances under which it will, in fact, be used." Lee (2015: 453) points out, "Various critics of Title II maintained that . . . Title II would be a non-transparent process and would not be administered according to a clear set of rules and settled precedents in sharp contrast to the Bankruptcy Code. These critics maintained that the [Senate] changes did not alter the fact that the federal government would be choosing which entities to resolve under Title II and which creditors to protect."

Brothers liquidation had catastrophic results for the firms' creditors. In the case of Lehman Brothers, senior bondholders received their first payment in April 2012, three and a half years after the bankruptcy filing. Unsecured creditors received their first distribution in September 2014, six years after the bankruptcy filing.¹³ Fleming and Sarkur (2014) found that, as of March 27, 2014, allowed claims to creditors stood at \$303.6 billion dollars. With its Thirteenth Plan Distribution to Senior Noteholders on October 5, 2017, Wilmington Trust reached a payout of approximately 42.7 percent to senior bondholders.¹⁴ Unsecured creditors have had a comparable payout percentage. Since the estate resolution is nearly complete, losses should end up at approximately 57 percent of the \$303.6 billion in allowed claims, or a loss of approximately \$173 billion to Lehman Brothers creditors. The time frame for payments to creditors, the scale of the losses, and the *ex ante* lack of clarity regarding how losses would be allocated to creditors could be disastrous for financial markets in future liquidations as they were with Lehman Brothers.

Conclusion

In this article, I have described a procedure that is capable of addressing the principle challenge of reorganizing failing financial institutions: maintaining the core intermediation and payment functions of the firm, avoiding a fire sale of its assets to cover liabilities, and allocating losses in a manner that is transparent and understood by a firm's creditors *ex ante*. Many of the challenges facing resolution regimes are obviated by our procedure. The firm remains as a going concern, maintaining both the core functions of the firm and all contractual obligations other than its long-term debt obligations to the holders of the reorganization bonds (which are replaced with equity). Concerns that have been raised by the FDIC regarding availability of debtor-in-possession financing are avoided, as are the incentives of foreign regulators and governments to ring-fence the assets of subsidiaries in their jurisdiction. Other concerns with resolution and liquidation regimes are also mitigated. Calabria (2015) points out that

¹³See Hughes Hubbard & Reed (2016: 27).

¹⁴See Wilmington Trust notices to senior noteholders (www.wilmingtontrust.com/lehman/notices.html).

“the Treasury . . . may have felt that allowing a default on GSE debt would be viewed internationally as the equivalent of a default by the U.S. government.” Concerns of this sort may lead regulators and politicians to ignore the law altogether and proceed with a bailout, as they did with the resolution procedures for Fannie Mae and Freddie Mac in 2008, or it may lead to pressures to include some favored creditors of a failing bank in the debts of the bridge bank and less fortunate or less favored creditors in the bankruptcy. The ambiguous status of creditors under both the Title I resolution plans and the Title II liquidation procedures leaves either of these possibilities open. Our procedure, by specifying particular long-term debt for conversion, lessens the pressure for a bailout and eliminates the possibility of favored treatment by government receivers for particular creditors of a failing financial institution.

The procedures could be developed in a new Chapter 14 of the bankruptcy code or in modifications to Chapter 11 for systemically important financial institutions. A great deal of work has been done to plan for the contingency that an important financial firm must enter bankruptcy or be liquidated. It would be good though to avoid those paths with a process that maintains all of the functions of a major financial institution without interruption, and prepositions liabilities that can be dedicated to recapitalize a failing financial firm in a manner that is known *ex ante* to regulators, to the firm’s creditors, and to other market participants.

References

- Calabria, M. (2015) “The Resolution of Systemically Important Financial Institutions: Lessons from Fannie and Freddie.” Cato Working Paper No. 25 (January 13). Available at https://object.cato.org/sites/cato.org/files/pubs/pdf/working-paper-25_1.pdf.
- Fleming, M. J., and Sarkur, A. (2014) “The Failure Resolution of Lehman Brothers.” Federal Reserve Bank of New York *Economic Policy Review* (December): 175–205.
- Gjerstad, S., and Smith, V. L. (2014) “Bonds, Not Bailouts, for Too Big to Fail Banks.” *Wall Street Journal* (August 12).
- Hamburger, P. (2014) *Is Administrative Law Unlawful?* Chicago: University of Chicago Press.
- Hughes Hubbard & Reed (2016) “State of the Estate” (August 16). *In re Lehman Brothers Inc.*, Case No. 08–01420 (SCC) SIPA.

- Lawson, G. (2015) “The Return of the King: The Unsavory Origins of Administrative Law.” *Texas Law Review* 93 (6): 1521–45.
- Lee, P. L. (2015) “Bankruptcy Alternatives to Title II of the Dodd-Frank Act—Part I.” *Banking Law Journal* (October): 437–87.
- McDermott, M. A. (2010) “Analysis of the Orderly Liquidation Authority, Title II of the Dodd-Frank Wall Street Reform and Consumer Protection Act.” Available at www.skadden.com/sites/default/files/publications/FSR_A_Analysis_Orderly_Liquidation_Authority.pdf.
- McDermott, M. A., and Turetsky, D. M. (2011) “Restructuring Large, Systemically-Important, Financial Companies: An Analysis of the Orderly Liquidation Authority: Title II of the Dodd-Frank Wall Street Reform and Consumer Protection Act.” *American Bankruptcy Institute Law Review* 19 (2): 401–51.
- Merrill, T. W., and Merrill, M. L. (2014) “Dodd-Frank Orderly Liquidation Authority: Too Big for the Constitution?” *University of Pennsylvania Law Review* 164: 165–247.

MONETARY POLICY IN AN UNCERTAIN WORLD: THE CASE FOR RULES

James A. Dorn

Since monetary policy operates in an uncertain world, discretionary policymaking relying on macroeconomic models of the economy is a weak reed upon which to base policy. The complexity of economic systems and constant changes in the underlying data mean errors may occur in a discretionary regime that can lead to monetary and financial instability.¹ The 2008 financial crisis is a case in point: central bankers and their expert staffs failed to anticipate the crisis, and may have worsened it by keeping policy rates too low for too long (Taylor 2012).

Moving to a rules-based regime would not eliminate radical uncertainty, but it could decrease institutional uncertainty—or what Robert Higgs (1997) has called “regime uncertainty”—and thus reduce the frequency of policy errors. Higgs focused on the uncertainty caused by fiscal and regulatory policies that attenuated private property rights by decreasing expected returns on capital.

Cato Journal, Vol. 38, No. 1 (Winter 2018). Copyright © Cato Institute. All rights reserved.

James A. Dorn is Vice President for Monetary Studies and a Senior Fellow at the Cato Institute. He thanks Ari Blask, Kevin Dowd, Thomas M. Humphrey, Gerald P. O'Driscoll, Alex Schibuola, and George Selgin for comments on earlier versions of this article. The usual disclaimer applies.

¹Orphanides (2017: 1) holds that “the presence of uncertainty . . . cannot serve as a valid argument for defending discretionary policy. Indeed, uncertainty raises the potential costs of discretion as it makes it harder to understand how large a policy deviation may be from what would have been the desirable systematic response to a shock.”

A discretionary monetary regime increases uncertainty about the future purchasing power of money and thereby undermines an important property right.

Radical uncertainty is a given, but institutional uncertainty can be reduced by adopting credible rules. As Karl Brunner (1980: 61) has pointed out:

We suffer neither under total ignorance nor do we enjoy full knowledge. Our life moves in a grey zone of partial knowledge and partial ignorance. More particularly, the products emerging from our professional work reveal a wide range of diffuse uncertainty about the detailed response structure of the economy. . . . A nonactivist [rules-based] regime emerges under the circumstances . . . as the safest strategy. It does not assure us that economic fluctuations will be avoided. But it will assure us that monetary policymaking does not impose additional uncertainties . . . on the market place.

In a similar vein, Allan Meltzer (1983: 95), a long-time collaborator with Brunner, has noted:

The flexibility that permits government to change policy has a cost: Anticipations about the future conduct of policy are altered. The effect of uncertainty is an important, but often neglected, characteristic that affects the cost of following alternative rules in a world subject to unpredictable changes.

Some congressional leaders think it's time to create a rules-based monetary regime. The Financial CHOICE Act of 2017 (H.R. 10), which recently passed the House, would make the Fed responsible for specifying a monetary rule and justifying to Congress any Fed deviations from it.²

Whether the CHOICE Act passes or not, it is important to consider alternative monetary rules and to be prepared to make the case for rules over discretion when the opportunity for reform arises.

This article begins with a discussion of the case for rules over discretion in the conduct of monetary policy and draws upon the theory of monetary disequilibrium to support that case. In particular, a

²See Financial CHOICE Act of 2017 (H.R. 10): https://financialservices.house.gov/uploadedfiles/hr_10_the_financial_choice_act.pdf.

credible monetary rule can eliminate what Clark Warburton (1949) called “erratic money,” which he viewed as the chief cause of business fluctuations.³

Various monetary rules will be examined, so will the difficulty of implementing them under the current environment in which unconventional Fed policy has plugged up the monetary transmission mechanism. Particular attention will be paid to rules designed to stabilize the path of nominal spending. The article ends with a call to establish a Centennial Monetary Commission to evaluate the Federal Reserve’s performance over its 100-plus years and to consider the effectiveness of alternative rules to reduce regime uncertainty.

The Case for Rules over Discretion

It is sometimes argued that discretionary monetary policy is superior to a rules-based monetary regime because discretion includes the option to adopt a rule. That argument, however, begs the question. The real issue is whether a robust, credible monetary rule that constrains policymakers to a long-run objective, and is strategic in nature, is superior to a discretionary regime that focuses on period-by-period optimization using various tactics without committing to any rule.

Those who favor discretion over rules also argue that no rule is permanent and thus judgment is needed to choose among rules. But choosing among rules is different from having no rule to guide policymakers, which is what is generally understood by a discretionary monetary regime. Under the Taylor rule, for example, one has to define the goal variables—the inflation gap and GDP gap—and use discretion in assigning numerical values to the coefficients on the goal variables. Nevertheless, it is still a rules-based monetary regime with a definite strategy as opposed to a regime that gives monetary authorities wide discretion (“the rule of experts”).

At the 2013 American Economic Association meeting, Lawrence H. Summers debated John Taylor on the issue of rules versus discretion. Summers used a medical analogy to make the case

³The analysis of alternative monetary rules in this article focuses on rules that are potentially stabilizing and omits rules such as the real bills doctrine that could be destabilizing. For an in-depth discussion of the real bills doctrine, see Humphrey and Timberlake (2017).

for discretion, arguing that he wants his doctor “to be responsive to the medical condition” rather than “to be consistently predictable.” Taylor responded by arguing that “relying on an all-knowing expert” who practices medicine without “a set of guidelines” is risky—“checklists are invaluable for preventing mistakes” just as a rules-based monetary strategy is. This argument in favor of rules is not to say that doctors don’t need to exercise good judgment in designing checklists. They do. But that kind of discretion needs to be distinguished from “a checklist-free medicine.”⁴ One could also argue that underlying Summers’s preference for “a doctor who most of the time didn’t tell me to take some stuff” is a fundamental rule: “Do no harm.”

Taylor (2015: 10) recognizes that “some rules are better than others, and it makes perfect sense for researchers and policymakers to be looking for new and better rules.” The focus should be on long-run strategy, not short-run tactics. The Fed did implicitly follow a Taylor rule during the Great Moderation, from the mid-1980s to 2007, and Taylor thinks that rule “does a good job at keeping nominal GDP on a steady growing trend.”

Taylor (2015: 4) does not recommend following “a rule mechanically”—“judgment is required to implement the rule.” He is thinking primarily of rules within the context of a government fiat money system. The question then becomes what is to bind policymakers to the rule. Although the Fed appears to have followed the Taylor rule in setting its policy rate during the Great Moderation, that adherence began to erode around 2003–05, when the fed funds rate was pushed significantly below the rate prescribed by the Taylor rule (*ibid.*, p. 5).

The Fed has not returned to any rules-based monetary policy even though Fed chairman Ben Bernanke argued in 2015 that the central bank was following a rule of “constrained discretion.” Yet, as Taylor observes, what Bernanke viewed as a rule—namely, setting goals (e.g., targeting inflation and employment)—differs substantially from adopting a rules-based monetary *strategy*. According to Taylor (2015: 12), “Simply having a specific numerical goal or objective function is not a rule for the instruments of policy; it is not a strategy; in my view, it ends up being all tactics.”

In order to better understand the case for rules over discretion, it is essential to recognize the knowledge problem confronting

⁴For a summary of the debate, see Taylor (2015: 11).

policymakers, the value of having time-consistent rules to reduce uncertainty, and the need to reduce the risk that monetary policy may become politicized as public choice theory describes.⁵

The Knowledge Problem

In his classic essay “The Use of Knowledge in Society,” F. A. Hayek (1945: 519–20) defined the “economic problem of society” as “a problem of the utilization of knowledge which is not given to anyone in its totality.” That problem implies monetary policymakers are not omniscient: they cannot know the structure of a complex economic system; their models will have serious flaws and forecast errors; there are long and variable lags in the effects of monetary policy, as noted by Milton Friedman (1968); and constant changes in economic data make it difficult to distinguish between permanent and transitory changes.

A discussion of the Hayekian knowledge problem, as it relates to monetary policy, is presented in O’Driscoll (2016). He argues that “unavoidable errors are an essential feature of discretionary policy” (p. 343), and that a rules-based monetary regime could help reduce uncertainty—an idea that both Hayek and Milton Friedman accepted. According to O’Driscoll (p. 350), “Hayek and Friedman agreed that we know too little to design an optimal monetary policy. . . . A monetary rule facilitates the emergence of a monetary order.”

Glenn Hubbard, former chairman of the Council of Economic Advisers under President George W. Bush, echoed those problems when he recently remarked, “Ignorance of economic conditions or doctrinaire attention to false models may blow Fed policy off course” (Hubbard 2017).

Nevertheless, Fed Vice Chairman Stanley Fischer, speaking at a Hoover Institution conference on May 5, 2017, argued that committees of experts rather than rigid rules are the best approach to sound monetary policymaking.⁶ According to Fischer, experts must “be continuously on the lookout for structural changes in the economy and for disturbances to the economy that come from hitherto unexpected sources.” However, the knowledge problem precludes such

⁵A useful overview of the case for rules over discretion is provided in Salter (2017: 444–48). See also Taylor (2017).

⁶For a critique of Fischer’s “rule-by-experts” approach to monetary policymaking, see White (2017).

changes and disturbances from being known beforehand; hence, Fed action is often destabilizing.

A discretionary monetary regime suffers most from these flaws and can be improved upon by moving to a rules-based regime (Friedman 1968). Monetary rules that are operational, credible, and enforceable could help reduce uncertainty.

Rules that are market based, don't rely on experts, and can evolve as learning occurs would be in line with Hayek's warning against the "pretense of knowledge." In his Nobel Memorial Lecture, Hayek ([1974] 1989: 7) stated: "To act on the belief that we possess the knowledge and the power which enable us to shape the processes of society entirely to our liking, knowledge which in fact we do not possess, is likely to make us do much harm." In monetary policy, relying on the Phillips curve and aggregate demand management to obtain full employment is an evident example of hubris.

The Fed's recent *Monetary Policy Report* (July 7, 2017) reflects the thinking of many Fed officials on the adoption of a money rule: "The U.S. economy is highly complex, and these rules, by their very nature, do not capture that complexity" (Board of Governors 2017: 36–37). In fact, it is the complexity of the economy that makes rules beneficial and more likely to bring about monetary and financial stability than pure Fed discretion.

Absence of Credible Commitment under a Discretionary Regime

Even if monetary authorities could centralize all the relevant information, a discretionary regime would not escape the problem of "time inconsistency" that Kydland and Prescott (1977) have pointed out.⁷ Under discretion, there is no guarantee that future policy will be consistent with current policy: monetary policymakers will be tempted, for example, to deviate from a commitment to maintain price stability in order to stimulate full employment by exploiting the short-run Phillips curve. Adherence to a monetary rule can improve policy outcomes if the rule amounts to a credible commitment.

⁷See White (1999: chap. 10) for an overview of the time-inconsistency case for rules, as first presented by Kydland and Prescott (1977) and elaborated on by Barro and Gordon (1983).

In his classic book *Interest and Prices*, Michael Woodford (2003) criticizes optimal control theory as applied to monetary policy while supporting the case for credible rules to reduce regime uncertainty and achieve long-run price stability. He opposes period-by-period policymaking, which he sees as destabilizing, compared to a rules-based regime:

It is not enough that a central bank have sound objectives . . . , that it make policy in a systematic way, using a correct model of the economy and a staff that is well-trained in numerical optimization, and that all this be explained thoroughly to the public. A bank that approaches its problem as one of optimization under *discretion*—deciding afresh on the best action in each decision cycle, with no commitment regarding future actions except that they will be the ones that seem best in whatever circumstances may arise—may obtain a substantially worse outcome, from the point of view of its own objectives, than one that commits itself to follow a properly chosen *rule* [Woodford 2003: 18–19].

Bennett McCallum, a member of the Shadow Open Market Committee, offers a similar criticism of discretionary policy:

The absence of rule-based policymaking means the absence of any systematic process that the public can understand and use as the basis for its expectations about future policy. The Fed apparently sees communication as a device for affecting expectations, but rational private agents form expectations on the basis of their understanding of the process by which the central bank actually conducts policy. If the central bank fails to adopt a process involving rule-based policymaking—that is, a commitment to some clearly stated objectives—its attempts to influence expectations are unlikely to be productive [McCallum 2004: 370].

A monetary rule is a constraint on the monetary authority in line with the rule of law. In addition to reducing regime uncertainty and increasing predictability of money and prices, a credible rule reduces the concentration of power over monetary matters and expands economic freedom. As Milton Friedman notes, in reference to the Fed's failure to maintain monetary stability during the Great Depression when the money supply contracted by nearly 30 percent, "much

harm can be done by mistakes on the part of a few men when they wield vast power over the monetary system of a country” (Friedman 1962: 50).

When policymakers have to follow known rules that recognize the limits of monetary policy and the social value of maintaining sound money, markets will be better able to perform their incentive, information, and allocation functions. There will be less risk of government intervention (e.g., price controls, credit allocation, and the politicization of investment) under a rules-based monetary regime than a discretionary regime. Government power grows, and economic freedom declines, when money and markets are in disorder, as we learned from the Great Depression, the high inflation and ensuing price controls of the 1970s, and the Great Recession, which greatly increased the Fed’s monetary and regulatory powers.

The Case for a Monetary Constitution

Monetary authorities have an incentive to increase their discretionary powers, especially during a crisis. When the Federal Reserve started operating in 1914, its powers were narrowly limited. Today, the Fed’s balance sheet has reached \$4.5 trillion and it engages in massive credit allocation and financial regulation. The 2008 financial crisis greatly expanded the Fed’s powers, and there is little incentive for Federal Reserve Board members to relinquish those powers.

Public choice theory describes how incentives operate within the public sector and how the administrative state feeds on itself. Constitutional political economy makes the case for limited government and the rule of law. To stem the incentives for monetary bureaucrats to expand their fiefdoms, their powers need to be limited by a strict monetary rule or constitution.

To be credible, a rule must be enforceable. In this regard, Selgin (2016b) distinguishes between “real and pseudo monetary rules.” In contrast to a pseudo monetary rule, which merely acts as a policy guide and is subject to change, a real monetary rule “must be rigorously enforced so that the public is convinced there will be no deviations from the rule.” It must also be “robust,” meaning “the rule must be capable of perpetuating itself, by not giving politicians or the public reason to regret its strict enforcement and to call either for its revision or its abandonment in favor of discretion” (p. 282).

The network of private contracts that characterized the classical gold standard, for example, were strictly enforced and allowed

markets to determine the quantity of money without interference from central banks. People had confidence in the long-run value of their money, which enabled them to borrow at reasonable interest rates for long periods of time. Commerce and investment were energized as a result.

While some economists (e.g., Buchanan 1962, 1989) favor a monetary constitution that retains a central bank but limits its powers, others (e.g., Hayek 1978) prefer a free-market monetary system. Between those two extremes there are many other sorts of monetary rules, many of which retain some degree of discretion for policymakers. For example, under inflation targeting, the monetary authority can arbitrarily change the target. Likewise, under the Taylor rule, more emphasis can be placed on reducing unemployment or closing the output gap rather than on achieving long-run price stability.

In making the case for rules over discretion, one should recognize that, under discretion, “the money-using public, uncertain about what the central bank experts will decide to do will hedge more and invest less in capital formation than they would with a credibly committed regime” (White 2017: 3). In contrast, “a commodity standard—especially without a central bank to undermine the redemption commitments of currency and deposit issuers—more completely removes policy uncertainty and with it overall uncertainty.” Consequently, there is a strong case for “a market-guided monetary system” rather than “expert-guided monetary policy” (*ibid.*).

The Theory of Monetary Disequilibrium

The case for a rules-based monetary regime can be better appreciated by an understanding of the theory of monetary disequilibrium—and thus the importance of limiting the power of central banks to manipulate money and interest rates.

Discrepancies between the quantity of money demanded and supplied at some prevailing price level set into motion an adjustment process during which real variables are influenced as the economy works its way toward a new equilibrium price level. Sticky prices and wages, as well as other institutional rigidities and the fact that money has no market of its own, mean not only that monetary disturbances are possible but also that they can have pervasive effects on real economic activity during the transition process.

According to Warburton (1949: 107), “The duration and amplitude of a business depression resulting from monetary disequilibrium

depends not only on the degree of that disequilibrium, but also on the tenaciousness of rigidities in the cost-price structure.” Thus, the theory of monetary disequilibrium may properly be called “a theory of the effect of price rigidities under an erratic supply of money.”⁸ Moreover, liquidity effects and other shorter-run real interest rate effects of monetary policy occur precisely because the price level doesn’t instantly adjust to disturbances in the demand and supply of money.

The fact monetary disequilibrium can persist for a significant time means that monetary policy can distort relative prices, especially intertemporal prices (i.e., interest rates) and misallocate capital. As Claudio Borio, head of the Monetary and Economic Department at the Bank for International Settlements, notes, “Monetary policy can fuel financial booms and their subsequent bust,” and in the process lead to long-lasting misallocation of resources, lower productivity, and a permanent loss of output (Borio 2016: 219–20).

Borio also addresses the idea that monetary disequilibrium can affect asset prices by driving a wedge between the market rate of interest and the equilibrium or natural rate that brings voluntary saving in line with private investment. According to Borio (2016: 214–19), in thinking about the Wicksellian natural rate of interest, it is not sufficient to consider only potential output and expected inflation; one must also consider financial stability. One cannot say that market rates are at equilibrium if there is financial instability. The natural rate of interest is unobservable; it is misleading to assume that the absence of inflation and the attainment of full employment signal that market rates are at their equilibrium levels—one must also check for a “build-up of financial imbalances.”

In Borio’s view, monetary policy

has failed to lean against unsustainable financial booms. The booms and, in particular, subsequent busts have caused long-term economic damage. Policy has responded very aggressively and, above all, persistently to the bust, sowing the seeds of the next problem. Over time, this has imparted a downward bias to interest rates and an upward one to debt [Borio 2016: 226].

⁸For a detailed discussion of the theory of monetary disequilibrium, see Warburton (1966, especially his list of postulates underlying that theory, pp. 28–29). Also see Yeager (1986, 1997) and Dorn (1987).

Central banks find it hard to increase interest rates because of fear that higher rates will deflate asset bubbles created by previous policy decisions to reduce interest rates. Moreover, politicians favor lower rates to keep the cost of deficit financing at bay. In such an environment, rates are likely to stay too low for too long, thus increasing the ultimate cost of adjustment.

It is also misleading, argues Borio (2016: 222–25), to view all episodes of deflation as bad. In particular, a gently falling price level during times of high real economic growth should be distinguished from sharply falling prices due to a prior monetary collapse, as happened from 1929–33, when the money supply fell by one-third during the Great Contraction (Friedman and Schwartz 1963). What should be avoided is demand-driven deflation, not productivity-induced deflation (Selgin 1997).

Finally, the idea that monetary policy can directly affect asset prices and distort investment decisions—even when official measures of inflation (such as the CPI or PCE) remain stable—is evident in the impact of recent, unconventional monetary policy on the prices of stocks, bonds, and real estate, with little impact on conventional measures of inflation.

Monetary disequilibrium theory holds that financial stability is best achieved by minimizing monetary instability, which means controlling the growth of money and credit to achieve stable growth of nominal GDP and long-run price stability. Under the classical gold standard, long-run price stability was ensured as the supply of money responded to changes in the demand for money. As Lawrence H. White (2017: 2) notes,

The actual track record of the classical gold standard is *superior* in major respects to that of the modern fiat-money alternative. Compared to fiat standards, classical gold standards kept inflation lower (indeed near zero), made the price level more predictable (deepening financial markets), involved *lower* gold-extraction costs . . . , and provided stronger fiscal discipline. The classical gold standard regime in the US (1879–1914), despite a weak banking system, *did no worse* on cyclical stability, unemployment, or real growth.

Central banks tried to improve on that regime but have ended up with a pure fiat money system not subject to any clearly defined monetary rule to reduce uncertainty about the future path of money and

prices—and business fluctuations in the United States have not lessened (Selgin, Lastrapes, and White 2012).

Alternative Monetary Rules

Monetary rules can be classified either as (1) limits placed on a discretionary central bank issuing government fiat money or (2) the replacement of a central bank with a market-based monetary system.⁹ Rules applicable to the first category include inflation targeting, a price-level rule, a Taylor rule, and demand rules aimed at achieving a stable growth path of nominal income. The second category of “rules” is greater in scope, but could, for example, consist of defining the dollar in terms of gold and allowing private banks to issue currencies convertible into gold.

Broad “meta-rules,” such as Peter Bernholz’s (2017: 100) call for “a concrete plan for a monetary constitution”—the key provision of which is “a constitutional safeguard that prevents governments and central bankers from influencing the stock of money”—are in line with Hayek’s call for “a constitution of liberty.”¹⁰ True meta-rules, such as the gold standard unconstrained by central banks, would be virtually devoid of discretionary elements. There would be no need to worry about defining and measuring policy objectives or estimating coefficients in equations representing the rule. Free markets, rather than policymakers, would operate to bring about monetary equilibrium.

Under Article 1, Section 8, of the U.S. Constitution, Congress has the authority “to coin Money [and] regulate the Value thereof.” Those enumerated powers have been delegated to the Federal Reserve, but with little oversight and wide discretion. As James Buchanan (1988: 33) has noted:

The dollar has absolutely no basis in any commodity base, no convertibility. What we have now is a monetary authority [the Fed] that essentially has a monopoly on the issue of fiat money, with no guidelines that amount to anything; an

⁹For an in-depth discussion of alternative monetary rules, see White (1999), Dorn (2017), and Salter (2017).

¹⁰For a discussion of meta-monetary rules, see Boettke, Salter, and Smith (2016). On the idea of a “monetary constitution,” see Yeager (1962) and White, Vanberg, and Köhler (2015).

authority that never would have been legislatively approved, that never would have been constitutionally approved, on any kind of rational calculus.

Indeed, today's pure fiat money system, and the lack of any monetary rule to limit discretion, is not something the Framers would have sanctioned. James Madison, the chief architect of the Constitution recognized that

the only adequate guarantee for the uniform and stable value of a paper currency is its convertibility into specie—the least fluctuating and the only universal currency. I am sensible that a value equal to that of specie may be given to paper or any other medium, by making a limited amount necessary for necessary purposes; but what is to ensure the inflexible adherence of the Legislative Ensurers to their own principles and purposes? [Madison 1831].

The courts and culture have eroded the Framers' monetary constitution (Timberlake 2013, Vieira 2017). Any link of the dollar to gold was officially ended in August 1971, when President Nixon closed the gold window. However, long before that event the Fed abandoned what Clark Warburton called “the convertibility theory of monetary control,” and never explicitly adopted the “responsibility theory of monetary control” to manage a fiat money regime. Under the convertibility theory, which was incorporated into the original Federal Reserve Act of 1913 but discarded by the monetary legislation of the early 1930s, the decisions of households and firms determine the quantity of money. The government's role is to ensure convertibility of notes and deposits into base (commodity) money. Under the responsibility theory, the decisions of central bankers determine the quantity of fiat money and maintain its value (Warburton 1966: 291–92).

The fact that present monetary law in the United States incorporates neither the *convertibility theory* nor the *responsibility theory* means monetary law remains in the same unsettled condition Warburton found it in 1946: “Monetary law in the United States is ambiguous and chaotic, does not contain a suitable principle for the exercise of the monetary power held by the Federal Reserve System, and has caused confusion in the development of Federal Reserve policy” (Warburton 1966: 316).

The Federal Reserve Reform Act of 1977 amended the Federal Reserve Act and implicitly adopted a monetary rule to limit growth in the monetary aggregates to long-run economic growth in order to achieve price stability. However, there was no operational rule and other objectives were added—namely, “maximum employment” and “moderate long-run interest rates.”¹¹ Moreover, there was no enforcement mechanism to hold the Fed responsible. Section 2A stated: “Nothing in this Act shall be interpreted to require that such ranges of growth or diminution [in the monetary aggregates] be achieved if the Board of Governors and the Federal Open Market Committee determine that they cannot or should not be achieved because of changing conditions.”

Today Section 2A simply reads:

The Board of Governors of the Federal Reserve System and the Federal Open Market Committee shall maintain long run growth of the monetary and credit aggregates commensurate with the economy’s long run potential to increase production, so as to promote effectively the goals of maximum employment, stable prices, and moderate long-term interest rates.

Yet the Fed pays little attention to monetary aggregates, has engaged in credit allocation to satisfy special interest groups, pays interest on excess reserves (IOER), and fails to recognize the limits of monetary policy in promoting long-run economic growth.

The Fed has rejected arguments for rules over discretion, showing that the Reform Act of 1977 did not succeed in truly implementing a monetary rule. Still, Congress can have the final say. Indeed, Congress is currently considering moving toward a rules-based regime and establishing a Centennial Monetary Commission to examine the Fed’s performance since its creation in 1913 and to consider various reforms.

The Financial CHOICE Act of 2017 makes the Fed responsible for specifying a monetary rule, while Congress would be required to use the Taylor rule as a benchmark or default rule. That rule specifies how the fed funds rate is to be adjusted based on the output gap (i.e., deviations of actual from potential output) and desired inflation, or more precisely, deviations of actual inflation from the chosen

¹¹See Pub. L. 95–188, 91 Stat. 1387, enacted November 16, 1977.

target (Taylor 1993). The Taylor rule rests on guesstimates of potential output and the Wicksellian natural rate of interest. The uncertainty regarding the value of those elements makes implementing the Taylor rule problematic, especially during times of financial turmoil. Furthermore, the rule formulates current policy based on past information as output and inflation measures are made available with lags.¹²

Simpler rules include: (1) Milton Friedman's (1960) k percent rule, which calls for money growth to be constant; (2) a price-level rule designed to achieve long-run price stability by controlling the monetary base; (3) inflation targeting; and (4) nominal income targeting. Friedman's k percent rule assumes that the demand for money (or its velocity) is stable. However, after the velocity of money became less stable, Friedman (1987) advocated freezing the monetary base and allowing the issuance of private bank notes.¹³

A price-level rule is plagued by long and variable lags in the relationship between money and prices,¹⁴ while an inflation-target rule may be destabilizing in the sense that a negative supply shock could temporarily increase inflation, leading to a tightening of monetary policy that would worsen the fall in output. For those reasons, there has been a resurgence of interest in demand rules aimed at stabilizing nominal GDP (NGDP).¹⁵

Scott Sumner, director of the Program on Monetary Policy at George Mason University's Mercatus Center, is a well-known proponent of NGDP targeting, which he thinks superior to alternative

¹²For these and other reasons, Beckworth and Hendrickson (2016) argue that the basic Taylor rule is inferior to a nominal GDP rule. Also see Selgin, Beckworth, and Bahadir (2015) on the case for a nominal GDP rule.

¹³One could argue that Friedman's k percent rule was never really tried and that if it had been, monetary velocity may have been more stable (see White 1999: 223).

¹⁴Haraf (1986: 361) has argued that, under a properly specified price-level rule, there would be increased certainty about future price levels that would improve the environment for nominal contracting. That improvement would reduce the lag between changes in the monetary base and the speed at which the observed price level approaches the target level. If so, a major objection to price-level targeting is removed. For further support of a price-level rule, see McCulloch (1991) and Dittman, Gavin, and Kydland (1999).

¹⁵Early proponents of nominal income targeting include Robert Hall (1981) and Robert Gordon (1985). George Selgin's "productivity norm" is also a type of demand rule, in which the price level would be allowed to vary inversely with real output while maintaining a stable path of aggregate spending (Selgin 1997).

monetary rules (Sumner 2014). One benefit of NGDP targeting is that it bypasses the issue of assigning weights under the Fed's dual mandate to achieve price-level stability and maximum employment. All that needs to be done is to set a target path for the growth of NGDP (i.e., the sum of real GDP growth and inflation). So, if the NGDP growth target is set at 5 percent, market forces will determine real growth and the Fed will supply the monetary base sufficient to hit the designated NGDP target. This strategy avoids having to fine tune monetary policy and would help circumvent the knowledge problem (see Beckworth 2017).

William Niskanen (1992: 284) has made a strong case for a demand rule targeting nominal domestic final sales. He argues that “a demand rule is superior to a price rule because it does not lead to adverse monetary policy in response to unexpected . . . changes in supply conditions. Similarly, a demand rule is superior to a money rule because it accommodates unexpected changes in the demand for money,” meaning unanticipated changes in the velocity of money. Niskanen sees base money as the best *instrument* to achieve a stable path for nominal income—and thus superior to using the fed funds rate.¹⁶

Bennett T. McCallum (1989: chap. 16) also calls for using the monetary base as an instrument. However, he prefers a feedback rule that, like Sumner's, “would aim at a zero inflation rate on average and would not attempt to be highly ambitious with regard to its effect on cyclical variation of real variables.” Accepting Warburton's (1949) argument against erratic money, McCallum seeks to avoid “abrupt changes in conditions due to monetary policy itself” (p. 338). He would allow the monetary base to grow in line with long-run real output growth adjusted for the growth in velocity averaged over the past four years. Nominal income would then tend to grow at a stable, non-inflationary rate reflecting the trend growth in real output.¹⁷

¹⁶Interest rates are not a good indicator of the stance of monetary policy: if the Fed increases money growth, and money incomes and inflation expectations rise, nominal interest rates will follow. Changes in base money are a better indicator, but only if base velocity is stable so there is a predictable relationship between base money, monetary aggregates, and nominal income. The best indicator is the behavior of spending itself.

¹⁷See White (1999: 223–24) for the simple analytics of the McCallum rule. Christensen (2011) provides “a market monetarist version of the McCallum rule.”

A Club of Financial Stability

Karl Brunner (1987: 49–51) once called for an international “club of financial stability” in which member states would agree to bind themselves to a monetary rule and thereby help reduce the uncertainty inherent in a discretionary government fiat money regime.¹⁸ Allan Meltzer (1989: 83) has argued that internal and external stability could be achieved if major countries each set “the rate of growth of the monetary base equal to the difference between the moving average of past real output growth and past growth in base velocity.” Doing so would anchor future expected prices and, with anticipated inflation stable, reduce the “variability of exchange rates arising from differences in expected rates of inflation” (ibid.).¹⁹

Meltzer’s rule is mildly activist but nondiscretionary; characteristics also present in McCallum’s (1984) rule. Choosing to stabilize the *anticipated* price level, rather than the actual price level, eliminates the need “to reverse all changes in the price level” (Meltzer 1989: 79). Instead, under Meltzer’s rule, the actual price level is allowed “to adjust as part of the process by which the economy adjusts real values to unanticipated supply shocks” (ibid.). In other words, Meltzer’s monetary rule “adjusts fully to permanent changes in growth rates of output and intermediation (or other changes in the growth rate of velocity) within the term chosen for the moving averages,” but ignores “short-term, transitory changes” (p. 81). Unlike a strict NGDP rule, Meltzer’s rule would accommodate persistent output changes with correspondingly more or less rapid money growth to achieve a mean-reverting long-run price level, much like that seen under the classical gold standard.

¹⁸On the importance of rules for obtaining monetary order and reducing the uncertainty present in a discretionary government fiat system, see Brunner (1985).

¹⁹Meltzer’s rule to stabilize the *anticipated* domestic price level of those countries who adopt his rule would still allow nominal exchange rates to vary with real exchange rates. In particular, “anticipated and actual exchange rates would be subject to change with changes in relative productivity growth, rates of growth of intermediation, differences in rates of saving, in expected returns to capital, in labor-leisure choice or other real changes” (Meltzer 1989: 80–81).

Unconventional Monetary Policy and the Plugged-Up Monetary Transmission Mechanism

In the current environment, with the Fed paying interest on excess reserves at a rate above what banks can get on highly liquid assets, the absence of a fully functioning fed funds market, and complex macroprudential regulations that discourage bank lending,²⁰ Meltzer's monetary rule (as well as other rules relying on the traditional links between base money, broader monetary aggregates, spending, and prices) would be difficult, if not impossible, to implement. In particular, by paying interest on excess reserves above the opportunity cost of those reserves, the Fed has increased the demand for holding excess reserves (rather than lending them out and creating a multiple expansion of deposits).²¹ Consequently, there has been a significant reduction in the size of the money multiplier, meaning there is a much weaker link between base growth and money growth than in the precrisis era (see Selgin 2017a, 2017b, 2017d).

Before serious consideration can be given to implementing any rule-based monetary regime, the Fed needs to normalize monetary policy by ending interest on excess reserves and shrinking its balance sheet to restore a precrisis fed funds market. Once changes in base money can be effectively transmitted to changes in the money supply and nominal income, the adoption of a monetary rule would reduce uncertainty and spur investment and growth.

Under the Fed's unconventional monetary policies, the growth rate of base money has far exceeded the growth of monetary aggregates and has not led to substantial, let alone rapid, growth of nominal income. Conventional price inflation has been tame.

²⁰On the impact of unconventional monetary policy on bank lending, especially the effect of overly zealous macroprudential regulation, see Calomiris (2017).

²¹There is no doubt that payment of interest on excess reserves (beginning in October 2008) at a rate exceeding interest on highly liquid assets (such as short-term Treasuries), has sterilized much of the newly created base money from the Fed's large-scale asset purchases. Humphrey (2014: 7) argues that paying IOER increases the "demand for idle reserves" and prevents them from being "lent out into active circulation in the form of bank deposit money." He notes that the Fed's attempt to expand the broad money supply to counter the financial crisis was hampered by paying IOER, which defeated the Fed's lender-of-last-resort function. According to Humphrey (in personal correspondence with the author), instead of paying a positive interest rate on excess reserves, the Fed should have charged a negative (penalty rate) to spur bank lending and deposit creation.

The large-scale purchases of longer-term Treasury securities and mortgage-backed securities have swelled the Fed's balance sheet to \$4.5 trillion from less than \$1 trillion before the crisis. However, banks have not lent out most of the new base money, and private investment has remained sluggish. Meanwhile, the Fed has used administrative measures to set a range for the fed funds rate—interest on excess reserves to set the upper limit and reverse repos to set the lower limit (see Selgin 2017a and Jordan 2016, 2017). The Fed has also engaged in credit allocation, used forward guidance to influence market perceptions of future rates, encouraged risk by underpricing it, penalized savers with ultra-low interest rates, and encouraged debt.

The problem is that, with massive amounts of excess reserves, there is no viable market for fed funds—more precisely, the only trading is arbitrage between GSEs that are not eligible for IOER and banks that are. Ending unconventional monetary policy by shrinking the Fed's balance sheet, while eliminating interest on excess reserves and the use of reverse repos, would help normalize monetary policy and restore the money multiplier to its precrisis values. The implementation of a demand rule would then be feasible.²²

Although the Fed has begun to increase the target range for the fed funds rate and has announced plans to shrink its balance sheet, the expectation is that the Fed will move very slowly and reverse course if asset prices tumble, disinflation occurs, or recession sets in.

Toward a Forecast-Free Monetary Regime

Leland Yeager (1992: 71) has proposed eliminating monetary disequilibrium by decentralizing and privatizing money, defining the unit of account “by a comprehensive bundle of goods and services,” and letting competition among private issuers “keep meaningful the denomination of their bank notes and deposits (and checks) in the stable, independently defined unit.”²³ He argues that those steps would take us much closer to a forecast-free monetary regime than our current government fiat money system under a highly

²²Belongia and Ireland (2015) have argued that the Fed could use Divisia monetary aggregates to make long-run targeting of NGDP feasible.

²³For a more detailed discussion, see Greenfield and Yeager (1983), and Yeager and Greenfield (1989).

discretionary central bank. Moreover, he is skeptical of “ideally managed government fiat money,” because that approach to monetary reform “precludes decentralizing and privatizing the issue of money” (*ibid.*). Absent a fundamental reform, he would favor a price-level rule over a demand rule.²⁴

Monetary Freedom and Monetary Order

A rules-based monetary system would increase economic freedom and lead to a more harmonious monetary order. The choice of what rule to follow will depend on whether one has more confidence in the convertibility theory of monetary control or the responsibility theory. There is no perfect monetary system, so tradeoffs must be made among competing rules. Furthermore, as digital currencies evolve, there may be completely novel ways to achieve monetary and financial stability.

Some critics of government fiat money believe that a gold standard could supply a desirable rule. Others would combine the properties of a gold standard with free banking or digital currencies. Still others would prefer binding the Fed by a monetary rule that is aimed at stabilizing NGDP, the price level, or inflation.

Lawrence H. White (2012) favors restoring “a gold definition of the U.S. dollar,” removing legal restrictions that prevent the emergence of a “parallel gold standard,” and allowing the issuance of private gold-backed currencies that could be used as legal tender. He provides a roadmap for making the transition to a new gold standard, but recognizes the difficulty of doing so without a broad public consensus. If that consensus does develop, however, financial innovation could help facilitate the transition.

Much of the criticism of monetary freedom has rested on the argument that free-market currencies are inherently unstable and inferior to a government-directed monetary system. Kevin Dowd (2017) constructs a hypothetical model of a *laissez-faire* monetary regime—asking how a free market in currencies would emerge

²⁴Bradley and Jansen (1989: 40) contend that changes in the assumptions about the labor market can make a price-level rule theoretically superior to a demand rule. Also, they argue that “ignorance of the correct equations, parameter values and lag structure that characterize the U.S. economy reduces the appeal of nominal GNP targeting.”

absent any central bank—and finds that its operating properties are consistent with stability and optimality, not chaos and inefficiency. The harmony that emerges under a market-based monetary system, argues Dowd, stems from the freedom to choose alternative currencies and the rule of law that binds the system together. White and Selgin discuss some historical examples of that stability.²⁵

Conclusion

Congress is currently considering moving toward a rules-based regime and establishing a Centennial Monetary Commission to examine the Fed's performance since its creation in 1913, and to consider various reforms. In doing so, it should not neglect the importance of restoring constitutional money and understanding how alternative monetary regimes affect uncertainty.

Normalizing monetary policy requires restoring the fed funds market and reducing the size of the Fed's balance sheet, which means eliminating interest on excess reserves and ending reverse repos so that selling longer-term Treasuries and mortgage-backed securities is accompanied by an equal reduction in excess reserves.

As Tatom (2017: 51) notes:

The Fed could repair its balance sheet and boost bank credit simply by reversing past actions. Since the last recession began, the Fed has accumulated about \$3.5 trillion of securities; 77 percent of bank receipts from these Fed purchases were added to excess reserves. Simply ending the subsidized interest on excess reserves would allow the Fed to sell the \$2.7 trillion of its securities held at the peak of excess reserves in August 2014 and incur a matching decline in banks' excess reserves. Such an operation would have no effect on the effective monetary base, monetary aggregates or total credit created in the money creation process. Fed credit and excess reserves would contract by \$2.7 trillion, but commercial bank credit would rise by an equal amount. This is precisely where risky assets should be held if banks are to promote growth and if the Fed is to get out of the credit allocation business.²⁶

²⁵See, e.g., White (1989), Selgin (2017c), Selgin and White (1987, 1994).

²⁶Selgin (2016a) provides a similar analysis.

The problem is that any reduction or even announcement of such a reduction in Fed assets could trigger a sharp fall in asset prices (especially in bond markets where duration risk is high) and shake market confidence—as seen in the 2013 Bernanke “taper tantrum” when he announced that the Fed would exit its large-scale asset purchase program. Financial markets have relied on the Fed “put” for a long time, and that expectation has made it difficult to change policy.

Also, unplugging the monetary transmission mechanism by ending interest on excess reserves, while desirable, would force the Fed to confront the problem of how to stop existing excess reserves from leaking out into the financial system and thereby creating inflation. The Fed may then face a period of stagflation and decide to revert back to unconventional monetary policy to “stimulate” the economy. Thus, the Fed is essentially in a trap that will be difficult to exit.²⁷ That is why it is essential to have a national debate over the direction of monetary policy and how best to reform the Fed.

References

- Barro, R. J., and Gordon, D. B. (1983) “A Positive Theory of Monetary Policy in a Natural Rate Model. *Journal of Political Economy* 91 (August): 589–610.
- Beckworth, D. (2017) “The Knowledge Problem in Monetary Policy: The Case for Nominal GDP Targeting.” Mercatus on Policy Series (July 18). Mercatus Center, George Mason University.
- Beckworth, D., and Hendrickson, J. R. (2016) “Nominal GDP Targeting and the Taylor Rule on an Even Playing Field.” Mercatus Working Paper, George Mason University (October).
- Belongia, M. T., and Ireland, P. N. (2015) “A ‘Working’ Solution to the Question of Nominal GDP Targeting.” *Macroeconomic Dynamics* 19 (3): 508–35.
- Bernanke, B. S. (2015) “Objections to Federal Reserve Accountability Bill.” Remarks presented at the conference on “The Fed in the 21st Century: Independence, Governance, and Accountability,” Brookings Institution, March 2. Available at www.youtube.com/watch?v=KJmA5JDNpKg.

²⁷On the unsustainability of unconventional monetary policies and the exit problem, see Dowd and Hutchinson (2017: 313–17).

- Bernholz, P. (2017) “The Implementation and Maintenance of a Monetary Constitution.” In J. A. Dorn (ed.) *Monetary Alternatives: Rethinking Government Fiat Money*, chap. 8. Washington: Cato Institute.
- Board of Governors, Federal Reserve System (2017) *Report on Monetary Policy* (July 7). Available at www.federalreserve.gov/monetarypolicy/files/20170707_mprfullreport.pdf.
- Boettke, P. J.; Salter, A. W.; and Smith, D. J. (2016) “Money as Meta-Rule: Buchanan’s Constitutional Economics as a Foundation for Monetary Stability.” George Mason University Working Paper in Economics, No. 16–49.
- Borio, C. (2016) “Revisiting Three Intellectual Pillars of Monetary Policy.” *Cato Journal* 36 (2): 213–38.
- Bradley, M. D., and Jansen, D. W. (1989) “Understanding Nominal GNP Targeting.” Federal Reserve Bank of St. Louis *Review* 71 (6): 31–40.
- Brunner, K. (1980) “The Control of Monetary Aggregates.” In *Controlling Monetary Aggregates III*, 1–65. Boston: Federal Reserve Bank of Boston.
- _____ (1985) “Monetary Policy and Monetary Order.” In *Monetary Policy and Monetary Regimes*, 4–21. Center Symposia Series No. CS-17. Rochester, N.Y.: Center for Research in Government Policy and Business, Graduate School of Management, University of Rochester.
- _____ (1987) “Policy Coordination and the Dollar.” Shadow Open Market Committee: Policy Statement and Position Papers (PPS 87-01). Center for Research in Government Policy & Business, University of Rochester.
- Buchanan, J. M. (1962) “Predictability: The Criterion of Monetary Constitutions.” In L. B. Yeager (ed.) *In Search of a Monetary Constitution*, 155–83. Cambridge, Mass.: Harvard University Press.
- _____ (1988) “Comment by Dr. Buchanan.” In *Prospects for a Monetary Constitution*, 32–35; special issue of *Economic Education Bulletin* 28 (6).
- _____ (1989) “Reductionist Reflections on the Monetary Constitution.” *Cato Journal* 9 (2): 295–99.
- Calomiris, C. W. (2017) “The Microeconomic Perils of Monetary Policy Experiments.” *Cato Journal* 37 (1): 1–15.

- Christensen, L. (2011) “A Market Monetarist Version of the McCallum Rule.” *The Market Monetarist*: <https://marketmonetarist.com/2011/10/09/a-market-monetarist-version-of-the-mccallum-rule>.
- Dittman, R.; Gavin, W. T.; and Kydland, F. E. (1999) “Price-Level Uncertainty and Inflation Targeting.” Federal Reserve Bank of St. Louis *Review* (July/August): 23–33.
- Dorn, J. A. (1987) “The Search for Stable Money: A Historical Perspective.” In J. A. Dorn and A. J. Schwartz (eds.) *The Search for Stable Money: Essays on Monetary Reform*, 1–28. Chicago: University of Chicago Press.
- _____, ed. (2017) *Monetary Alternatives: Rethinking Government Fiat Money*. Washington: Cato Institute.
- Dowd, K. (2017) “Monetary Freedom and Monetary Stability.” In J. A. Dorn (ed.) *Monetary Alternatives: Rethinking Government Fiat Money*, chap. 18. Washington: Cato Institute.
- Dowd, K., and Hutchinson (2017) “From Excess Stimulus to Monetary Mayhem.” *Cato Journal* 37 (2): 303–28.
- Fischer, S. (2017) “Committee Decisions and Monetary Policy Rules.” Speech presented at “The Structural Foundations of Monetary Policy,” a Hoover Institution Monetary Policy Conference, Stanford University (May 5).
- Friedman, M. (1960) *A Program for Monetary Stability*. New York: Fordham University Press.
- _____. (1962) *Capitalism and Freedom*. Chicago: University of Chicago Press.
- _____. (1968) “The Role of Monetary Policy.” *American Economic Review* 58 (1): 1–17.
- _____. (1987) “Monetary Policy: Tactics versus Strategy.” In J. A. Dorn and A. J. Schwartz (eds.) *The Search for Stable Money*, 361–82. Chicago: University of Chicago Press.
- Friedman, M., and Schwartz, A. J. (1963) *A Monetary History of the United States, 1867–1960*. Princeton, N. J.: Princeton University Press.
- Gordon, R. J. (1985) “The Conduct of Domestic Monetary Policy.” In A. Ando et al. (eds.) *Monetary Policy in Our Times*. Cambridge, Mass.: MIT Press.
- Greenfield, R. L., and Yeager, L. B. (1983) “A Laissez-Faire Approach to Monetary Stability.” *Journal of Money, Credit, and Banking* 15 (August): 302–15.

- Hall, R. (1981) “Lowering Inflation and Stimulating Economic Growth.” In *Politics and the Oval Office: Toward Presidential Governance*, 207–27. San Francisco: Institute for Contemporary Studies.
- Haraf, W. S. (1986) “Monetary Velocity and Monetary Rules.” *Cato Journal* 6 (2): 641–62.
- Hayek, F. A. (1945) “The Use of Knowledge in Society.” *American Economic Review* 35 (4): 519–30.
- _____ ([1974] 1989) “The Pretense of Knowledge: Nobel Memorial Lecture of December 4, 1974.” *American Economic Review* 79 (6): 3–7.
- _____ (1978) *The Denationalisation of Money*. 2nd ed. London: Institute of Economic Affairs.
- Higgs, R. (1997) “Regime Uncertainty: Why the Great Depression Lasted So Long and Why Prosperity Resumed after the War.” *The Independent Review* 1 (4): 561–90.
- Hubbard, G. (2017) “How to Keep the Fed from Following Its Models Off a Cliff.” *Wall Street Journal* (June 15).
- Humphrey, T. M. (2014) “Averting Financial Crises: Advice from Classical Economists.” Federal Reserve Bank of Richmond *Economic Focus* 18 (4): 4–7.
- Humphrey, T. M., and Timberlake, R. H. (2017) “The Real Bills Doctrine, the Gold Standard, and the Great Depression.” Unpublished manuscript.
- Jordan, J. (2016) “Tools of Monetary Policy.” *Sound Money Project* (December 13): <http://soundmoneyproject.org/2016/12/tools-of-monetary-policy>. A project of the Atlas Network.
- _____ (2017) “Rethinking the Monetary Transmission Mechanism.” *Cato Journal* 37 (2): 361–84.
- Kydland, F. E., and Prescott, E. C. (1977) “Rules Rather than Discretion: The Inconsistency of Optimal Plans.” *Journal of Political Economy* 85 (June): 473–91.
- Madison, J. (1831) “[Letter] to Mr. Teachle” (Montpelier, March 15). In S. K. Padover (ed.), *The Complete Madison: His Basic Writings*, 292. New York: Harper and Bros. (1953).
- McCallum, B. T. (1984) “Monetary Rules in the Light of Recent Experience.” *American Economic Review* 74 (May): 388–96.
- _____ (1989) *Monetary Economics: Theory and Policy*. New York: Macmillan.

- _____ (2004) “Misconceptions Regarding Rules vs. Discretion for Monetary Policy.” *Cato Journal* 23 (3): 365–72.
- McCulloch, J. H. (1991) “An Error-Correction Mechanism for Long-Run Price Stability.” *Journal of Money, Credit, and Banking* 23 (August, Part 2): 619–24.
- Meltzer, A. H. (1983) “Monetary Reform in an Uncertain Environment.” *Cato Journal* 3 (1): 93–112. Reprinted in J. A. Dorn and A. J. Schwartz (eds.) *The Search for Stable Money: Essays on Monetary Reform*, 201–20. Chicago: University of Chicago Press, 1987.
- _____ (1989) “On Monetary Stability and Monetary Reform.” In J. A. Dorn and W. A. Niskanen (eds.) *Dollars, Deficits, and Trade*, 63–85. Boston: Kluwer. This paper was originally presented at the Third International Conference of the Institute for Monetary and Economic Studies at the Bank of Japan, June 3, 1987.
- Niskanen, W. A. (1992) “Political Guidance on Monetary Policy.” *Cato Journal* 12 (1): 281–86.
- O’Driscoll, G. P. Jr. (2016) “Monetary Policy and the Knowledge Problem.” *Cato Journal* 36 (2): 337–52.
- Orphanides, A. (2017) “The Case Against the Case Against Monetary Rules.” Paper presented at the Shadow Open Market Committee meeting, Princeton Club, New York City (May 5).
- Salter, A. W. (2017) “Some Political Economy of Monetary Rules.” *The Independent Review* 21 (3): 443–64.
- Selgin, G. (1997) *Less than Zero: The Case for a Falling Price Level in a Growing Economy*. Hobart Paper No. 132. London: Institute of Economic Affairs.
- _____ (2016a) “Interest on Excess Reserves, Part 2.” *Alt-M* (January 5): www.alt-m.org/2016/01/05/interest-reserves-part-ii.
- _____ (2016b) “Real and Pseudo Monetary Rules.” *Cato Journal* 36 (2): 279–96.
- _____ (2017a) “On Shrinking the Fed’s Balance Sheet.” *Alt-M* (February 23): www.alt-m.org/2017/02/23/shrinking-the-feds-balance-sheet.
- _____ (2017b) “A Monetary Policy Primer, Part 10: Discretion or a Rule?” *Alt-M* (May 11): www.alt-m.org/2017/05/11/a-monetary-policy-primer-part-10-discretion-or-a-rule.
- _____ (2017c) *Money: Free and Unfree*. Washington: Cato Institute.

- _____ (2017d) “Monetary Policy v. Fiscal Policy: Risks to Price Stability and the Economy.” Testimony before the U.S. House of Representatives Committee on Financial Services (July 21). Available at www.cato.org/publications/testimony/monetary-policy-v-fiscal-policy-risks-price-stability-economy.
- Selgin, G.; Beckworth, D.; and Bahadir, B. (2015) “The Productivity Gap: Monetary Policy, the Subprime Boom, and the Post-2001 Productivity Surge.” *Journal of Policy Modeling* 37 (2): 189–207.
- Selgin, G.; Lastrapes, W. D.; and White, L. H. (2012) “Has the Fed Been a Failure?” *Journal of Macroeconomics* 34 (3): 569–96.
- Selgin, G., and White, L. H. (1987) “The Evolution of a Free Banking System.” *Economic Inquiry* 25 (July): 439–57.
- _____ (1994) “How Would the Invisible Hand Handle Money?” *Journal of Economic Literature* 32 (December): 1718–49.
- Sumner, S. B. (2014) “Nominal GDP Targeting: A Simple Rule to Improve Fed Performance.” *Cato Journal* 34 (2): 315–37.
- Tatom, J. A. (2017) “How to Fix the Fed: The Ineffectiveness of the U.S. Central Bank’s Credit Policy.” *The International Economy* (Winter): 48–51.
- Taylor, J. B. (1993) “Discretion versus Policy Rules in Practice.” *Carnegie-Rochester Conference Series on Public Policy* 39: 195–214.
- _____ (2012) “Monetary Policy Rules Work and Discretion Doesn’t: A Tale of Two Eras.” *Journal of Money, Credit, and Banking* 44 (6): 1017–32.
- _____ (2015) “Getting Back to a Rules-Based Monetary Strategy.” Written version of keynote address given at the Shadow Open Market Committee’s Conference on “Getting Monetary Policy Back on Track,” Princeton Club, New York City, March 20.
- _____ (2017) “Rules versus Discretion: Assessing the Debate over the Conduct of Monetary Policy.” Paper presented at the Federal Reserve Bank of Boston Conference on “Are Rules Made to Be Broken? Discretion and Monetary Policy” (October 13).
- Timberlake, R. T. (2013) *Constitutional Money: A Review of the Supreme Court’s Monetary Decisions*. New York: Cambridge University Press.

- Vieira, E. Jr. (2017) “Gold and Silver as Constitutional Alternative Currencies.” In J. A. Dorn (ed.) *Monetary Alternatives: Rethinking Government Fiat Money*. Washington: Cato Institute.
- Warburton, C. (1949) “Erratic Money: An Outline of the Theory of Monetary Disequilibrium.” Box 16, Warburton Collection, Special Collections, George Mason University Library, Fairfax, Va. (Unpublished book-length manuscript.)
- _____ (1966) *Depression, Inflation, and Monetary Policy: Selected Papers, 1945–1953*. Baltimore: The Johns Hopkins University Press.
- White, L. H. (1989) *Competition and Currency: Essays on Free Banking and Money*. New York: New York University Press.
- _____ (1999) *The Theory of Monetary Institutions*. Oxford: Basil Blackwell.
- _____ (2012) “Making the Transition to a New Gold Standard.” *Cato Journal* 32 (2): 411–21.
- _____ (2017) “Experts and the Gold Standard.” *Alt-M* (June 13). Available at www.alt-m.org/2017/06/13/experts-gold-standard.
- White, L. H.; Vanberg, V. J.; and Köhler, E. A., eds. (2015) *Renewing the Search for a Monetary Constitution*. Washington: Cato Institute.
- Woodford, M. (2003) *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton, N.J.: Princeton University Press.
- Yeager, L. B., ed. (1962) *In Search of a Monetary Constitution*. Cambridge, Mass.: Harvard University Press.
- _____ (1986) “The Significance of Monetary Equilibrium.” *Cato Journal* 6 (2): 369–99.
- _____ (1992) “Toward Forecast-Free Monetary Institutions.” *Cato Journal* 12 (1) 53–73.
- _____ (1997) *The Fluttering Veil: Essays on Monetary Disequilibrium*. Edited by G. Selgin. Indianapolis: Liberty Fund.
- Yeager, L. B., and Greenfield, R. L. (1989) “Can Monetary Disequilibrium Be Eliminated?” *Cato Journal* 9 (2): 405–21.

REFORMING THE RULES THAT GOVERN THE FED

Charles W. Calomiris

As historians of the Fed such as Allan Meltzer (2003, 2009a, 2009b, 2014) frequently note, the Fed has failed to achieve its central objectives—price stability and financial stability—during about three-quarters of its first 100 years of operation. Although the Fed was founded primarily to stabilize the panic-plagued U.S. banking system, since the Fed’s founding the United States has continued to suffer an unusually high frequency of severe banking crises, including during the 1920s, the 1930s, the 1980s, and the 2000s. Unfortunately, research shows that the Fed has played an active role in producing most of those crises, and its failure to maintain financial stability has often been related to its failure to maintain price stability. The Fed-engineered deflation of the 1930s was the primary cause of the banking crises of that era.

Cato Journal, Vol. 38, No. 1 (Winter 2018). Copyright © Cato Institute. All rights reserved.

Charles W. Calomiris is Henry Kaufman Professor of Financial Institutions at Columbia University, Director of the Program on Financial Studies at Columbia Business School, Research Associate of the National Bureau of Economic Research, Member of the Shadow Open Market Committee, Member of the Financial Economists Roundtable, Co-Director of the Hoover Institution Initiative on Regulation and the Rule of Law, and a Fellow at the Manhattan Institute. An earlier version of this article was presented as testimony on April 4, 2017, before the U.S. House of Representatives Subcommittee on Monetary Policy and Trade of the Committee on Financial Services. This article draws upon prior work, including a coauthored 2015 Shadow Open Market Committee (SOMC) presentation to congressional staff written by Charles Calomiris, Greg Hess, and Athanasios Orphanides. Allan Meltzer, Mickey Levy, John Taylor, Athanasios Orphanides, and Deborah Lucas provided helpful comments.

The Fed's lax monetary policy produced the Great Inflation of the 1960s and 1970s, which was at the heart of the interest rate spikes and losses in real estate, agricultural, and energy loans during the 1980s, which produced the banking crisis of that period. A combination of accommodative monetary policy from 2002 to 2005, alongside Fed complicity with the debasement of mortgage underwriting standards during the mortgage boom of the 2000s, and Fed failures to enforce adequate prudential regulatory standards, produced the crisis of 2007–09 (Calomiris and Haber 2014: chs. 6–8).

It is worth emphasizing that the U.S. experience with financial crises is not the global norm; according to the IMF's database on severe banking crises, the two major U.S. banking crises since 1980 place our country within the top quintile of risky banking systems in the world—a distinction it shares with countries such as Argentina, Chad, and the Democratic Republic of Congo (Laeven and Valencia 2013, Calomiris and Haber 2014).

In his review of Fed history, Allan Meltzer (2003, 2009a, 2009b, 2014) points to two types of deficiencies that have been primarily responsible for the Fed's falling short of its objectives: (1) adherence to bad ideas (especially its susceptibility to intellectual fads); and (2) politicization, which has led it to purposely stray from proper objectives. Failures to achieve price stability and financial stability reflected a combination of those two deficiencies.

Unfortunately, the failures of the Fed are not merely a matter of history. Since the crisis of 2007–09, a feckless Fed has displayed an opaque and discretionary approach to monetary policy in which its stated objectives are redefined without reference to any systematic framework that could explain those changes, has utilized untested and questionable policy tools with uncertain effect, has been willing to pursue protracted fiscal (as distinct from monetary) policy actions, has grown and maintains an unprecedentedly large balance sheet that now includes a substantial fraction of the U.S. mortgage market, has been making highly inaccurate near-term economic growth forecasts for many years, and has become more subject to political influence than it has been at any time since the 1970s. The same problems that Meltzer pointed to—bad ideas and politicization—now, as before, are driving Fed policy errors. I am very concerned that these Fed errors may result, once again, in departures from price stability and financial stability (Calomiris 2017a, 2017b, 2017c).

In this article, I show that the continuing susceptibility of the Fed to bad thinking and politicization reflects deeper structural problems that need to be addressed. Reforms are needed in the Fed's internal governance, in its process for formulating and communicating its policies, and in delineating the range of activities in which it is involved. I will focus on three types of reforms that address those problems: (1) internal governance reforms that focus on the structure and operation of the Fed (which would decentralize power within the Fed and promote diversity of thinking); (2) policy process reforms that narrow the Fed's *primary* mandate to price stability and that require the Fed to adopt and to disclose a systematic approach to monetary policy (which would promote transparency and accountability of the Fed, thereby making its actions wiser, clearer, and more independent); and (3) other reforms that would constrain Fed asset holdings and activities to avoid Fed involvement in actions that conflict with its monetary policy mission (which would improve monetary policy and preserve Fed independence).

Together these three sets of reforms would address the two most important recurring threats to monetary policy—short-term political pressures and susceptibility to bad ideas—and thereby improve the Fed's ability to achieve its ultimate long-run goals of price stability and financial stability, which are crucial to promoting full employment and economic growth. Table 1 summarizes the reforms proposed here, and Figure 1 outlines the primary channels through which reforms would improve monetary policy.

The Need for Internal Governance Reforms

The Fed needs broad and fundamental changes to its internal governance. Internal governance reform should make the Fed more institutionally democratic and more diverse in its thinking. Those improvements, in turn, would make the Fed less susceptible to political pressure—because centralization of power *in* the chair invites more political pressures *on* the chair. They also would make the Fed less likely to adhere to bad ideas, because of a reduced likelihood of “group think.” My proposed changes are unlikely to have strong internal advocates within the Fed system (at the very least inside the beltway, at the Board of Governors), and therefore will require legislation. Ironically, although the Fed has been a

TABLE 1
SUMMARY OF PROPOSED REFORMS

Internal Governance Reforms to Fed Structure and Operation

1. Require at least two of seven Fed governors to be people with significant financial markets experience.
 2. Governors should each have at least two staff members under their direct control.
 3. Require governors to resign other positions as a condition for appointment.
 4. Require governors to pledge that they expect to stay for at least half of their appointed terms.
 5. Increase salaries for governors to ensure that the Fed remains able to attract talented people.
 6. Enhance retirement benefits for governors, contingent on a sufficient number of years of service.
 7. Federal Reserve Bank presidents should be selected by their Board of Directors, not subject to the approval of the Board of Governors. At the very least, adopt a sunshine law that requires the Board of Governors to provide summary information to Congress regarding any candidates the Federal Reserve Board rejects, as well as information about candidates that the Board suggests for consideration, or asks to be dropped from consideration prior to being formally proposed by the Banks.
 8. All Reserve Bank presidents should vote at every FOMC meeting.
 9. Budgetary authority should rest in a committee comprising representatives of all the Federal Reserve Banks and the Board of Governors, and perhaps even some outsiders, to determine the budgets of each Bank and each governor's staff.
 10. Prohibit Reserve Bank presidents from being promoted from within their own Bank.
-

Policy Process Reforms

11. Replace the "dual mandate" with a single price-stability primary objective.
 12. Require the Fed to maintain and state a systematic approach to monetary policy. The policy framework would be controlled by the Fed, and subject to change as the Fed sees fit.
-

(Continued)

TABLE 1 (*Continued*)
SUMMARY OF PROPOSED REFORMS

Avoiding Inappropriate Policies or Conflicts of Interest

13. Prohibit the Fed from holding securities other than U.S. Treasury securities in its portfolio (except during emergencies, in the context of assistance approved under its emergency lending powers).
 14. Interest on reserves should be set at 10 basis points below the federal funds rate.
 15. Prohibit the Fed from engaging in its current reverse repos as a substitute for open market operations.
 16. Remove the Fed from writing and enforcing regulations. The Fed would still participate in examinations and have full access to all information necessary to fulfill its role as a lender of last resort. At the very least, the Fed should be removed from merger decisions and oversight of highly politically sensitive matters, such as CRA examinations.
-

U.S. Budgetary Reform

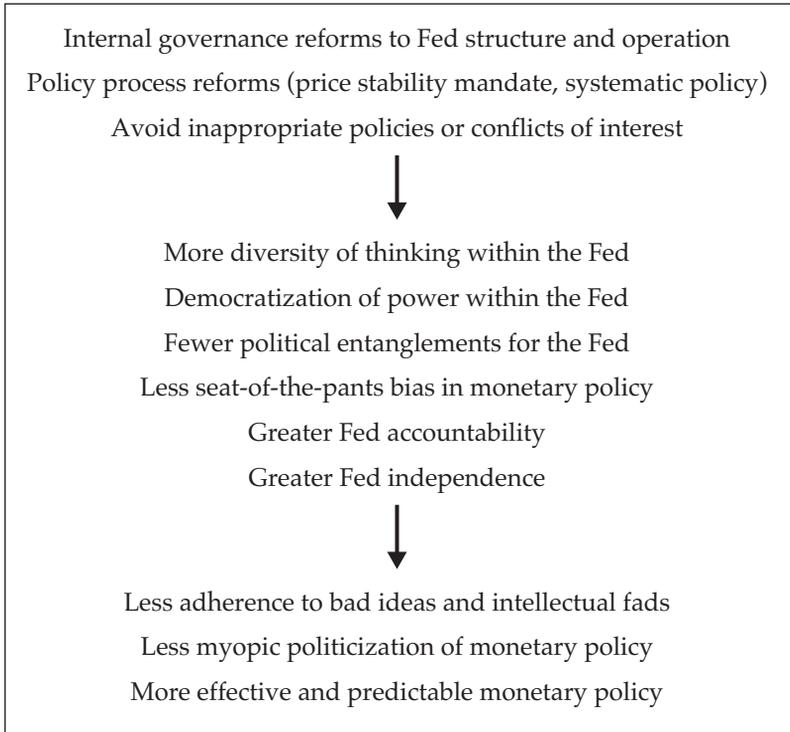
17. The Fed's surplus revenues should not be used as an off-budget means of funding the Consumer Financial Protection Bureau (CFPB), other regulatory actions, highway expenditures (including those undertaken by the Fed), or other programs.
-

champion of governance reform for banks as a means of improving their performance, it is much less receptive to recognizing its own governance problems.

In recent years, there has been an unhealthy increase in the centralization of power within the Fed, which has two parts: (1) the power of the Fed chair over the Federal Reserve Board; and (2) the concentration of power within the Federal Reserve System at the Board of Governors.

Daniel Thornton and David Wheelock (2014), both economists who have served for many years at the Federal Reserve Bank of St. Louis, provide some heuristic evidence on the need to reduce the power of the Fed chair over the Board of Governors. Thornton

FIGURE 1
HOW PROPOSED GOVERNANCE REFORMS WOULD
IMPROVE MONETARY POLICY



and Wheelock report that Federal Reserve Board governors have dissented from the chair only two times from 1995 to 2014. This compares to 65 dissents during the same period of time by Federal Reserve Bank presidents. Interestingly, presidents and governors had a similar pattern of dissents from 1957–95, about eight dissents per year for each group.

Surely, a well-informed and diligent group of six independent (nonchair) governors would find reason to disagree from time to time with the chair. Federal Reserve Bank presidents dissent frequently, and Supreme Court justices dissent with aplomb. Dissents remain common at the Bank of England. But somehow, Fed governors in recent years have become restrained from expressing their dissenting views.

This lack of dissent would seem strange to architects of the current Fed structure. When then-Fed chair Marriner Eccles testified before the Senate Banking Committee on March 4, 1935, regarding the proposed structure of the Federal Open Market Committee (FOMC), he complained that including only three Federal Reserve Board governors ran the risk that “a minority of the Board [of Governors] could adopt a policy that would be opposed to one favored by the majority [of the seven board members].” That argument convinced Congress to structure the FOMC to include all seven governors. Clearly, Eccles envisioned a healthy degree of potential dissent within the Board of Governors about monetary policy. That is no longer the case.

Three possible explanations emerge for this unhealthy trend toward uniformity at the Board of Governors, each of which indicates a need for reform. One possibility is that governors are selected based on their willingness to compromise and “to go along, to get along.” The chair has substantial discretionary power that can be wielded against uncooperative governors. This possibility, if true, is indicative of an unhealthy internal governance system that quashes independent thinking.

A second possibility is that many of the governors have become, de facto, short-timers who may not have a permanent stake in the system’s long-run management or performance. Why bother to dissent if you are leaving soon after arriving? If this explanation has merit, it indicates that Fed governors are not playing the role intended by the Federal Reserve Act, which entrusted them with significant authority, gave them long-term (14-year) appointments, and envisioned them as important contributors to shaping the policies of the Fed.

Finally, a third possibility is that governors may not have the information or background needed to support the formation of independent decisions. This is quite possible given that (nonchair) governors do not have *any* staff to support their own lines of research and inquiry, and historically their access to the Board’s staff has been limited. To the extent that limits are sometimes relaxed by the chair, this is a discretionary decision that can be reversed (and perhaps *would* be reversed if governors made use of staff to support positions that opposed the chair). Fed governors have complained publicly about the lack of independent staff to advise them, or the inability to speak to staff without permission. Former vice chairman Alan Blinder

frequently complained about the limitations placed on his ability to communicate with Fed staff, and also complained about the “real reluctance to advance alternative points of view” at the Fed. Former governor Laurence Meyer said that he was “frustrated by the disproportionate power the chairman wielded over the FOMC,” and said that dissents were viewed as disruptive to the process of monetary policymaking (Calomiris 2014).

Not only is there a disturbing power imbalance within the Board of Governors, but also there has been a shift in power within the System toward Washington. Throughout the founding and operation of the Fed, it has always been recognized that the Board of Governors is more attuned to political pressures than the Fed presidents. The presidents, therefore, play a crucial role in deterring political influences that tend to make monetary policy myopic. The shift of power toward the Board has made it harder for Federal Reserve Bank presidents to challenge the point of view coming from the chair, and serves to politicize the Fed (e.g., through pressures applied by the administration to the chair).

One symptom of the shift of power toward Washington has been an increasingly aggressive “approval” process by the Board of Governors for nominees to be presidents of Federal Reserve Banks. The Fed Board of Governors has approval authority over the appointment of presidents, but recently they have been taking a more aggressive role in suggesting nominees and refusing to approve others. Although the discussion of this issue has been limited to those within the system—as well as journalists—recent presidential searches have purportedly resulted in the nonapproval of multiple finalist candidates put before the Federal Reserve Board by Boards of Directors of the Federal Reserve Banks.

In addition to the problems of excessive centralization of power in Washington and in the Fed chair, there has been a cultural shift at the Fed that has reduced the diversity of thinking and made the Fed more susceptible to academic fads. As recently as the 1970s, Fed governors and presidents typically were not academics steeped in the latest modeling fads of macroeconomics. But in recent years, it has become rare for governors or presidents to be people coming from backgrounds other than academia. This likely reflects several influences, including the increasing centralization of power within the system noted above, and changes in the structure of the banking

industry; but it also probably reflects the increasing technical complexity of macroeconomic modeling, which many non-PhD economists find challenging to comprehend.¹

The models the Fed has employed for policy purposes, however, have not proven to be of much value. The fashionable “dynamic stochastic general equilibrium” (DSGE) model, which was all the rage in Fed and academic thinking during the years leading up to the crisis, conceived of the economy as divorced from the banking sector (a sector that was not important enough to be included in the DSGE model). Needless to say, the banking crisis proved that this was an important omission. Since the 2007–09 crisis, DSGE models have been modified to try to incorporate the financial sector. Nevertheless, the consistent failure of the Fed to forecast economic growth over the past decade gives little reason for confidence in current Fed modeling.

In the past, many of the most successful Fed leaders did not put much stock in the latest fads of macroeconomic modeling. It is widely believed that Paul Volcker was the most successful Fed chair of the past several decades. In one Fed cartoon prepared for high school students, Paul Volcker is lovingly portrayed as a superhero wearing a red cape. Few would object to that characterization. Mr. Volcker’s combination of integrity, judgment, and courage stand alone:

- Integrity because—prior to his appointment—he leveled with President Carter about his intention to attack inflation aggressively.
- Judgment because he rejected the model-driven advice of some top Fed economists who adhered to “Phillips curve”-based projections; Volcker recognized that only a draconian policy change would be sufficient to establish Fed credibility in lowering inflation.

¹The rise of nationwide branch banking in the 1990s caused important local and regional banks to largely disappear, which has changed the profiles of Federal Reserve Banks’ boards. The increasing rigor of Fed modeling at FOMC meetings (despite the inaccuracy of that modeling, especially in the years leading up to the subprime crisis) has fostered a culture that makes it quite difficult for nonacademics to challenge the assumptions of the chair’s preferred econometric model, however misspecified it may be. Even someone like Alan Greenspan, a trained economist who worked outside of academia and who resisted placing too much weight on forecasts from the Fed’s macroeconomic models, is missing in the ranks of Fed leadership today.

- Courage because he stayed the course despite sustained high unemployment and vilification from many academics who derided his policies—because they contradicted the received academic wisdom of the day.

If the Fed were to face a similar challenge again—and the risks associated with its balance sheet’s size and structure make that a real possibility—would someone emerge with the same combination of virtues? Sadly, the answer is perhaps not. People like Volcker—who took macroeconomic modeling with the appropriately large grain of salt, whose spine was stiffened by years in the trenches of global banking, and who deeply understood the psychology of financial markets—are unlikely to end up as leaders of today’s Fed.

That fact would not have pleased the Fed’s founders. The structure of the System, as originally conceived, and as reformed in 1935, was designed to ensure a healthy diversity of experience among its leaders. Fed leadership was supposed to combine those with experience in banking with political appointees with different life experience. Academics were absent from leadership positions, as they were not selected as political appointees until much later—Arthur Burns was the first academic to serve as chair. A system of 12 Federal Reserve Banks was intended to ensure that Fed leaders would be guided by diverse regional *banking* perspectives. Even at the Board, banking professionals sometimes dominated (e.g., Marriner Eccles was a Utah banker, and Paul Volcker worked at Chase when he wasn’t at the Treasury or the Fed).

Some Fed leaders I have spoken with tell me that nonacademics often lack understanding of key economic issues. That may be true, but every governor or president doesn’t have to understand statistics deeply to be able to contribute to the collective wisdom of the Fed. Sometimes the most important contribution one can make in a meeting is to question things that economists as a group accept too easily. It is worth emphasizing that group think about models has been a perennial problem at the Fed. In the 1920s–60s, it was the Riefiler-Burgess “net free reserves” model; later it was the Phillips Curve; and more recently, the New Keynesian DSGE model.

Don’t get me wrong: technical modeling is necessary, but it is not helpful to fill the FOMC with people who use the same model. We need multiple models, and we need people who bring other facts and

thinking to bear on economic questions. There is no better antidote to Fed group think than having FOMC members who are willing to scoff at economists' certainties, especially if their own experiences provide credible alternative perspectives about how markets and people behave.

Promoting Democratization of Power and Diversity of Thinking within the Fed System

It is possible to construct new rules for Fed leadership that will reduce the centralization of power, enhance diversity, and reduce group-think risk.

Congress could require, for example, that at least two of the seven Fed governors be people with significant financial markets experience. Having at least two people on the Board with backgrounds in the financial industry—like Peter Fisher and Kevin Warsh—would create a critical mass of market-savvy opinion.

To further build diverse thinking at the Board, the power of the chair should be limited. For starters, to ensure that governors have access to necessary information and can act independently in their voting, governors should each have at least two staff members under their direct control, which would enable them to develop independent views.

Perhaps that reform would help to solve another problem: the short tenure of most governors. Governors' terms are 14 years, but most leave after only two years. Before governors become fully educated to the intricacies and challenges of monetary policy, they are on their way back to the universities whence they came (to avoid losing their chaired professorships). Congress should require governors to resign their other positions, including university professorships, as a condition for appointment, and also ask them to pledge that they expect to stay on as governors for at least half of their appointed terms. Salaries for governors should also be increased to ensure that the Fed remains able to attract talented people. Part of the reason that governors return to academia so quickly is that for most of them their salaries as governors are much less than what they earn at universities. Furthermore, after two years on the FOMC, lucrative consulting and private board of directors appointments beckon. Retirement benefits for governors could also be enhanced, and made contingent on a sufficient number of years of service.

In addition to reforming the Board of Governors, Congress should increase the role of Federal Reserve Banks within the FOMC and increase their independence within the Fed System. Federal Reserve Bank presidents should be selected based on the independent decisions of their Board of Directors, and should not be subject to the approval of the Board of Governors. At the very least, if the Board of Governors is to retain its approval power, let's adopt a sunshine law that requires it to provide summary information to Congress (which maintains all candidates' privacy) regarding any candidates the Federal Reserve Board rejects, as well as information about candidates that the Board suggests for consideration, or asks to be dropped from consideration prior to being formally proposed by the Banks.

Congress also should change FOMC voting rules so that all Reserve Bank presidents vote at every meeting. That would promote diversity by giving more power and voice to the research staffs of the Reserve Banks. Current FOMC rules of rotation are designed to give greater weight to the Board, which effectively means the huge research staff controlled by the Fed chair.

Perhaps most important, the 12 Federal Reserve Banks should also be freed from the budgetary control of the Fed Board and its chair, who can use budgetary power (e.g., to limit the size and scope of their research activities) as a threat to gain cooperation on policy matters. For example, the Federal Reserve System could establish a committee comprising representatives of all the Federal Reserve Banks and the Board of Governors, and perhaps even some outsiders, to consider the budgets of each Bank and each governor's staff.

It would further promote diversity of thinking if Federal Reserve Banks were prohibited from appointing Reserve Bank presidents from within their own Bank. When Federal Reserve Banks' Boards were composed of regional banking and business leaders, Boards had a direct stake in Fed decisionmaking and presidents were selected from a broad pool of outsiders. Now, almost all Fed presidents are former Bank research economists (usually research directors). Although formal searches are always undertaken, it is hard to attract qualified outsiders to participate in that process when they know that the internal candidate has the inside track, based on his or her relationship with the Board, and even if they do participate, risk-averse Boards often prefer the devil they know to the one they don't. The result is unhealthy inbreeding.

Policy Process Reforms: A Primary Price-Stability Mandate and Systematic Policy

The internal governance reforms outlined above must be supplemented with policy process reforms that ensure the right kind of accountability for the Fed by improving policy transparency, constraining unaccountable discretion, and discouraging politicization of monetary policy. Fed history shows that some of the Fed's worst errors were the result of the wrong kind of accountability. As Allan Meltzer's (2003, 2009a, 2009b, 2014) work shows, including his three volume *History of the Federal Reserve*, Fed failures often have reflected political pressures on the Fed to accommodate deficits, or an excessive focus on short-term unemployment goals (with an eye to upcoming elections) at the expense of long-term inflation and unemployment goals. An important safeguard against monetary policy errors, therefore, is to promote greater independence of the Fed.

Paradoxically, unlimited Fed discretion does not result in greater independence of action because unlimited discretion invites political interference. Fed independence is best achieved by imposing discipline on the process of monetary policy in a way that sets clear objectives for policy and enhances accountability with respect to achieving those objectives.

The most obvious policy process improvement would be to repeal the "dual mandate" imposed on the Fed in the 1970s and replace it with a single primary price-stability mandate, as Paul Volcker and Alan Greenspan, among many others, have publicly championed.² The sole primary objective of price stability is embodied in many other central banks' charters, including those of the European Central Bank and the Bank of England. There are three arguments for adopting this policy in the United States.

First, price stability is an achievable long-run objective, and thus the Fed can be held accountable for achieving it. Indeed, long-run inflation is *completely* under its control. The Fed has a monopoly over the supply of currency. Inflation is the (inverse of the) value

²With respect to the financial stability mandate, focusing monetary policy on price stability would also tend to avoid financial instability. Of course, aside from monetary policy, there are other important regulatory policy tools that should be used to promote financial stability (see Calomiris 2017a).

of money; if you control its supply, you control its value. Unpredictable short-term changes in demand and measurement problems make this hard to do on a short-term basis, but over sufficient time the Fed can control inflation. In contrast, the Fed cannot be held accountable for achieving a given unemployment target in the long run; indeed, economists agree that the long-run rate of unemployment is the result of factors outside of the control of the Fed.

Second, inflation matters for growth. High levels of inflation, or volatile inflation, result in lower output and higher unemployment in the long run. As Milton Friedman and many others correctly argued for many years, the reason to target price stability is not that we care about price stability per se (no one should), but rather because we care about employment and output; by making price stability the primary long-run objective of the Fed we ensure that *the average levels of output and employment will be maximized in the long run*. Paradoxically, the point of narrowing the Fed's long-term mandate to inflation is to boost average employment.

Third, narrowing the Fed's primary mandate protects it from myopic political pressures that are inherent in any democracy. Elections can lead politicians to pressure the monetary authority to make the wrong tradeoffs, such as boosting output today (in the interest of current voters) at the expense of higher inflation and lower output tomorrow (at the expense of future voters). Giving the Fed a narrow price stability primary objective provides cover for the Fed in defending itself against opportunistic attacks. Complicating monetary policy by introducing multiple goals (unemployment alongside price stability) makes it hard to hold the Fed accountable for its actions in the long run, while encouraging political pressures on the Fed to achieve an amorphous employment objective. I believe that the Fed's risky QE3 program of purchasing mortgage-backed securities (MBS) and long-term Treasury bonds in an effort to demonstrate its commitment to reducing unemployment (which had very little effect in boosting employment) is an example, among many, of how such myopic political pressures can distort monetary policy.

The call for a single price-stability mandate is often misunderstood as reflecting a callous lack of interest in unemployment, but the opposite is the case. Economic studies have shown that in the long run there is no tradeoff between price stability and maximum employment; it follows that a single primary long-run commitment to

price stability in no way requires a tradeoff of lower employment. Holding the Fed primarily to account for price stability does not preclude it from supporting the economy during slumps with countercyclical policy over the short or medium terms, as a *secondary* objective. Indeed, the Taylor Rule is an example of a policy that is consistent with both meeting a long-run inflation target and providing countercyclical influence. There is no doubt that a Fed with a single inflation mandate would continue to execute countercyclical policy aggressively. However, making price stability its sole primary objective ensures maximum *sustainable* growth and employment over the long run, while defending the independence of the Fed from short-term political pressures.

In addition to narrowing the Fed's primary mandate to price stability, it would also enhance accountability and independence to require the Fed to maintain a *systematic approach* to monetary policy. This is crucial for two reasons: First, systematic policy defends against the dangers of discretionary, seat-of-the-pants policymaking that is susceptible to biases and socio-political pressures. One of the most important seat-of-the-pants biases is "dynamic inconsistency." As academics and Fed researchers have long recognized, a non-rule-based monetary policy will tend to err both with respect to hitting its inflation target and with respect to optimally stabilizing the economy over the business cycle. These twin seat-of-the-pants biases are sometimes referred to as "inflation bias" and "stabilization bias" (see, for example, Faust and Henderson 2004), and are part of a long tradition in monetary policy research emphasizing the social welfare improvements that come from adherence to long-term commitments by the monetary policy authority (e.g., Friedman 1948, 1959, 1968; Phelps 1968, 1972; Kydland and Prescott 1977; Orphanides 2003a, 2003b). Systematic rule-based policy is key to avoiding undue focus on the short term at a long-term cost to society.

Second, because businesses and households take decisions that depend on expectations about the future, including future policy decisions and their economic consequences, a systematic policy framework that makes monetary policy more predictable facilitates better private-sector decisions over time and enhances social welfare.

What constitutes systematic monetary policy? Clearly, not the status quo, which delegates monetary policy to Fed policymakers with broad mandates and allows these policymakers to employ unconstrained judgment in meeting those mandates.

Over its first century of operation, however, the Fed sometimes acted relatively systematically. Variation in the quality of Fed policymaking over time reflects, in part, variation in the degree to which policy was systematic and oriented toward clear long-term objectives. One reason for this diversity in outcomes over different periods is the immense discretionary power that the Federal Reserve has exercised over time in interpreting its mandate and in deciding monetary policy, as well as the lack of any effective oversight of the monetary policy process. Recognizing that appointed policymakers are humans and are susceptible to all the pressures and biases that humans face, reform legislation can play a crucial role in giving monetary policy a clear long-term focus and forcing it to implement a systematic policy process. Policy so conceived would be more predictable and less susceptible to fads, to short-term seat-of-the-pants biases, and to myopic political influences.

Does a systematic approach to monetary policy imply adherence to a rigid, static rule? If economists had a perfect understanding of the economy and the ability to observe and properly interpret shocks, and if the structure of the economy were unchanging, then monetary policy could be guided by a fixed policy rule that would specify how the Fed would react to observed shocks to maintain price stability and smooth the business cycle. But that is not a realistic vision of what systematic monetary policy would mean in the real world of changing economic structure and imperfect economic understanding.

Economists always have an incomplete understanding of the economy and face limitations in observing and interpreting shocks hitting the economy in real time, when policy decisions counteracting potential adverse effects of various shocks have to be made. As a result, there are divergent views and considerable uncertainty regarding precisely what the best monetary policy response to macroeconomic conditions may be. Furthermore, as the structure of the economy evolves over time, any algebraic rule characterizing the appropriate policy response to macroeconomic factors will have to adapt to changing circumstances. In addition, policymakers learn over time, and their understanding of appropriate policy responses, therefore, is also subject to change even if the structure of the economy is not changing.

The limitations introduced by these sources of uncertainty have been used by some to justify relying on policymakers' "best judgment"

with unlimited discretion. That is a fatuous argument. So long as the systematic formulation of monetary policy is flexible and able to adapt to changes in the economy's structure and our understanding, uncertainty cannot justify the resistance to making policy systematic.

This point bears emphasis: the adoption of a simple, but flexible, monetary policy rule is clearly desirable because it can tackle uncertainty about the economy while avoiding the adverse consequences of unlimited discretion.

Let me be clear: a policy rule must be a specific algebraic formula that can be used to determine how monetary policy should respond to changes in macroeconomic conditions, as summarized by specific observable variables such as the current inflation and unemployment rates. By its very nature, such a policy rule ensures that policy is systematic, transparent, and accountable. If well designed (based on existing empirical evidence), the policy rule will also deliver good economic performance. Research with estimated models of the U.S. economy over the past few decades suggests that simple policy rules can be designed that would deliver good economic performance. Of course, there is reasonable disagreement about what the best rule would be, and care is required both in the evaluation of alternative policy rules and in their implementation. Sifting through this evidence and reaching appropriate judgments about which rule to apply, and adapting the rule over time as needed, should be the central functions of any monetary policy authority.

Taylor (2016) discusses recent progress in the evaluation of macroeconomic models, like those that would form the basis for the evolving Fed policy rule. Two important characteristics of the model evaluation process are noteworthy. First, models must be developed in what Taylor calls the "rules space" rather than the "path space." Models in path space conceive of policy as the execution of isolated, hypothetical one-time policy actions. Models in rules space evaluate alternative policy rules in a framework in which policy actions occur within the context of the rules that produce them. Not only are rules-space models the only coherent approach to model the effects of policy on the economy (because, for example, they take account of expectations that are influenced by the existence of the rule), they are also ideally suited to inform an FOMC that is charged with developing and constantly improving its explicit monetary rule. Second, the model evaluation process must identify common performance criteria that would be used to evaluate the relative validity of a

diverse range of models. Volker Wieland's pioneering efforts to develop "The Macroeconomic Model Data Base" (see www.macro-modelbase.com) shows that this is possible (see Wieland et al. 2016). Wieland's website invites all comers to propose models, and provides a platform in which they can be compared, debated, and verified empirically. An FOMC charged to follow and disclose its systematic approach to monetary policy could benefit from making use of just such a website. And a Fed structured to encourage diverse thinking would make effective use of it.

One might ask whether a flexible policy rule could be an effective constraint on unbridled discretion. After all, the FOMC would be free to change its rule at every meeting. Yes, it would, but it would have to do so as a committee, reaching agreement on the changes needed, and embodying those beliefs in observable parameter changes that outsiders could challenge. Outside opinions about the quality of FOMC deliberations and decisions about its rule would be a source of accountability, including at Congressional hearings, and the FOMC would have reason to care about its reputation as a crafter of empirically defensible rules.

To help fix ideas about how FOMC discussions would be likely to proceed, consider an example policy rule for the federal funds rate, f , based on the well-known Taylor (1993) rule:

$$(1) \quad f = r^* + \pi + a(\pi - 2) - b(u - u^*)$$

This rule suggests that when the inflation rate equals the 2 percent target and the unemployment rate equals the natural rate of unemployment, u^* , monetary policy should be neutral, that is, the federal funds rate should be equal to the sum of the real natural rate of interest, r^* , and the inflation target. If inflation is above the target, then policy should be tighter, with the degree of the policy response depending on the parameter a (in Taylor's original formulation this was equal to 0.5). If the unemployment rate is above the natural rate of unemployment, as is typically observed in recessions, monetary policy should be eased, with the response governed by the parameter b . Considering different values for the parameters a and b is a simple way to see that care is required to ensure that a policy rule, if followed, will contribute to good economic outcomes over time. If b is set to zero, this rule does not respond at all to employment conditions and may lead to undesirable volatility in unemployment. If b is set to a very high value, say 10, this rule becomes very activist and

may result in undesirable instability in both inflation and unemployment. Choosing parameters that would deliver the best macroeconomic performance depends on one's beliefs about the economy. Alternative estimated models typically suggest somewhat different values for the parameters that work best.

The Taylor-type rule above also highlights an important issue relating to the natural rate of unemployment and the natural rate of interest. These concepts are unobservable and are typically estimated. However, estimates are uncertain and may vary considerably both over time and due to differences in estimation methodologies. Using a Taylor-type rule with the wrong estimates of the natural rates introduces a bias that results in deviations from price stability. As an example, the original formulation of the Taylor rule was based on the assumption that the natural rate of interest is 2 percent. Some analysts, including Federal Reserve officials, presently suggest that their preferred current estimates of the natural rate of interest are zero or even negative. The same policy rule with these two alternative assumptions would give policy prescriptions from the Taylor rule that would differ by 200 basis points or more.

I emphasize, however, that such disagreements about hard-to-measure concepts like the natural rate of interest are not insurmountable obstacles to agreeing on a rule. Indeed, alternative policy rules could be specified that do not depend on estimates of the natural rates to set policy and are therefore not subject to related uncertainty. For example, a rule might employ the values of the federal funds rate and the unemployment rate in the previous quarter, f_{-1} and u_{-1} .

$$(2) \quad f = f_{-1} + a (\pi - 2) - b (u - u_{-1})$$

Compared to the Taylor rule, this rule suggests that the federal funds rate should be raised (relative to its value a quarter earlier) if inflation is above the target and should be reduced if the unemployment rate is higher than it was in the previous quarter.

Simple rules, based on the examples above, can also make use of forecasts of economic activity and inflation. Indeed, there are many reasonable candidates for a simple policy rule. A critical issue in determining which rule the monetary authority should adopt among the many alternatives is how robust the rule is to the various sources of uncertainty and potential error. In a committee setting, such as the FOMC, there may be differences of opinion about what

is the appropriate way to think about the U.S. economy that may not be possible to distinguish on the basis of available empirical evidence.

A reasonable criterion for designing a simple rule for the Federal Reserve would be the robustness of the rule to reasonable alternative models. This is how policy ought to be designed to defend against major inference errors in an environment of uncertainty.

Requiring the Fed to identify and adopt a policy rule along the lines highlighted above would replace meeting-by-meeting discretion and thus ensure that the harmful consequences of seat-of-the-pants policy are avoided. But as I have noted, given the complexity and continuous change of the economy, it would not be expected that any simple algebraic formula could be the basis for robust policymaking forever.

The goal in making monetary policy systematic is not to replace discretionary policy with an immutable rule, but rather to replace it with a systematic framework for selecting a simple and robust rule that foresees periodic reviews and adaptation. Nor would this process of discussing and disclosing come as an unprecedented innovation within the Fed. Publication of the simple rule that the FOMC would follow has a precedent in the Fed's current publication of principles regarding its longer-run goals, which the FOMC has been publishing every January since 2012. As the FOMC adapts its rule over time, to ensure that best practices prevail in the evaluation process, it would be important that a high degree of transparency accompany the process of evaluation of alternative rules and any adaptations under consideration.

Crucially, the selected rule should be specified with sufficient detail to hold the FOMC accountable and eliminate meeting-by-meeting discretion. An outside observer should be able to determine the meeting-by-meeting setting of policy *using only public information*. If the rule's implementation requires use of unobserved concepts that may vary from quarter to quarter, such as the natural rate of interest, then the methodology for tracking those changes over time should be made explicit so that it could be replicated with public information. Similarly, if the rule employs short-term projections of inflation, these projections should be in line with those available to the public. In other words, unaccountable discretion should not be introduced through the back door, for example by using a simple rule that responds to inflation projections based on policymakers' "judgment" that cannot be independently reproduced and evaluated.

Because no simple rule can encompass satisfactorily crisis situations that might require a rapid policy response, an escape clause should be included that allows policy to deviate from the simple rule. In the past few decades, a few instances could be identified, perhaps once every decade or two, when a deviation from a simple rule could be necessary. To cover such contingencies, a comply-or-explain approach should be adopted, with the understanding that deviations are rare and related explicitly to crisis circumstances.

Providing the Fed with a single primary mandate of price stability and requiring it to maintain a systematic, flexible approach to policy are reforms that are long overdue. Several senior Fed policymakers recently mischaracterized the current legislative proposal (U.S. House of Representatives 2016) to require monetary policy to be systematic as dictating an immutable rule, such as the Taylor Rule, to the Fed. This is disingenuous. My understanding of the current proposed legislation is that it conforms to the proposal I lay out here: the Fed would determine its own policy rule, which would be subject to its decisions to alter the rule over time, and in emergency circumstances the Fed would not be rigidly bound to adhere to its stated framework.

The methodology and expertise necessary for the Federal Reserve to adopt a simple and robust policy rule that can preserve price stability and deliver good stabilization performance are available. Requiring a systematic approach will have a constructive effect on the substance of FOMC deliberations and their information content for outside observers, by encouraging much of the debate to focus on whether to revise the existing framework, and how to do so. This will make monetary policy more predictable, more understandable to the market, more accountable to Congress, and more independent of myopic political pressure. Given the undisputed benefits of avoiding seat-of-the-pants policymaking, preserving the Fed's unlimited discretionary approach cannot be reasonably defended.

Limits on Fed Activities and Holdings

There are major problems that arise from combining monetary policy with other functions and powers. Most obviously, a systematic rule for monetary policy may mean little if it is only one of many things that the monetary policy authority is doing. Unaccountable discretion could arrive through the back door of other policies and

undermine the commitment to systematic policy. And those other policies, because they would not be subject to the discipline of systematic thinking and accountability, would invite myopic political influence. Furthermore, other mandates on the Fed related to regulatory policy, or its own financial interests, may conflict with its role as a monetary policy authority.

These are not hypothetical problems. The Fed's current fiscal and regulatory policy actions give it many policy levers other than those related to traditional monetary policy. Without reforms that limit Fed actions and holdings, even if the internal governance and policy process reforms suggested above were implemented, the Fed would continue to suffer from conflicts of interest and politicization risk that could encourage it to choose inferior monetary policy rules, or to undermine the effects of its systematic monetary policy rule with other actions. Additional reforms, therefore, are needed to avoid the conflicts and politicization that result from the current multiple roles, powers, and instruments of the Fed.

The Fed's powers and toolkit have grown since the crisis of 2007–09. One of the most remarkable aspects of Dodd-Frank was the confidence it evinced in the Fed. The Office of Thrift Supervision was abolished after the 2007–08 crisis in response to its perceived incompetence. But Dodd-Frank enhanced the supervisory and regulatory powers of the Fed (which was a primary regulator of several of the most deeply troubled banks, including Citi and Wachovia).

That enhancement of Fed power was all the more remarkable when one considers that in March 2008, the U.S. Treasury circulated a “blueprint” explaining why it would be desirable to redesign the U.S. financial regulatory structure along functional lines. That change also would have reduced the conflicts of interest inherent in exercising of monetary policy and regulatory authority by removing many supervisory and regulatory powers from the Fed (Calomiris 2006, 2013). Under the “blueprint,” the Fed would continue playing a key role in examinations, with full access to information that might be useful to it in its capacity as lender of last resort, but it would not play a central role in the rule setting or supervision of banks. The “blueprint” was put aside after the crisis, which largely reflected the skill of Fed advocates (especially Chairman Bernanke) in convincing Congress that the Fed was the most able and trustworthy party in which to vest many of the new regulatory powers created by Dodd-Frank.

Since the crisis, as the Fed's powers have grown, so have its conflicts of interest. In particular, monetary policy experimentation has involved the Fed as a direct participant in financial markets in unprecedented ways. As of February 22, 2017, the Fed holds \$1.8 trillion dollars in MBS on its balance sheet (which amounts to roughly one-sixth of the U.S. mortgage market), reflecting the Fed's new role in spurring the economy by subsidizing mortgage finance costs. It is noteworthy that this new fiscal policy role of the Fed was not primarily the result of crisis support, but rather of Fed purchases of mortgage-backed securities as part of its "quantitative easing" experiments.

Many critics regard this as an inappropriate incursion into fiscal policy by the Fed. It also creates numerous conflicts of interest with respect to the Fed's role as a regulator of banks. As a holder of MBS, the Fed has an incentive to avoid actions that might increase mortgage interest rates, even if that would be desirable as a matter of monetary policy. This is true for two reasons. First, any accounting losses on its MBS portfolio would increase the Fed's contribution to the measured deficit, with obvious adverse political ramifications.³ Second, housing finance is a magnet for political interests, which implies severe continuing pressures on the Fed not to sell its mortgage portfolio, even if failing to do so serves to prop up a destabilizing housing bubble.

The Fed also sets interest rates banks earn on reserves. The Fed apparently intends to use this tool to potentially offer very high interest rates, if necessary, to dissuade banks from lending, as an alternative to selling its portfolio (and recognizing politically unappealing capital losses when doing so). Although Title II, Sec. 201, of the Financial Services Regulatory Relief Act of 2006, which authorized the payment of interest on reserves, clearly limited Fed discretion in setting interest payments by specifying that "balances maintained at a Federal Reserve Bank . . . may receive earnings . . . at a rate or rates not to exceed the general level of short-term interest rates," the Fed appears to intend to side-step this legislative limit by creating a

³According to the Fed's accounting rules, the Fed does not mark its portfolio to the market; it incurs losses on securities only if those securities are sold. The Fed's capital losses affect the measured deficit, but on a consolidated basis they have no economic effect on the government's deficit. Nevertheless, they matter politically, as critics of the Fed are likely to make use of its measured contribution to the deficit. Because that threat is real, the Fed will seek to avoid sales of assets that cause its measured contribution to the deficit to rise.

“range” of targeted values, with the interest on reserves expected to lie at the top of the specified range, and by reserving its right to “adjust the interest on excess reserves rate . . . as necessary for appropriate monetary control, based on policymakers’ assessments of the efficacy and costs of their tools.”⁴ Some politicians have already challenged the Fed to explain why it is appropriate for it to pay above-market rates on bank reserves. Clearly, this is a fiscal expenditure, just as paying zero interest is the commonly understood “reserve tax.” It is inappropriate for the Fed to make fiscal decisions about the taxes or subsidies transferred from the government to banks, and it is doubly inappropriate, given the Fed’s role as a bank regulator.

Not only do the Fed’s holdings of MBS and its setting of interest on reserves entail new fiscal actions and politicization risks, but also the Fed now acts as a repo counterparty, and will do so increasingly over time. This new activity (like setting interest on reserves) appeals to the Fed because it provides the Fed a means for avoiding the politically embarrassing recognition of capital losses that it would otherwise incur if it sold long-duration securities into the market as interest rates rise. Rather than sell securities from its portfolio to contract its balance sheet, the Fed engages in reverse repos, repeatedly lending those securities into the market until they mature, and thus avoiding sale while effectively reducing its balance sheet size.

Over the past several decades, repo has been an important alternative source of funding for lending in the U.S. economy, by both regulated banks and nonbank lenders. The massive expansion of the Fed’s balance sheet over the past decade has withdrawn a large amount of low-risk collateral from the market, thereby making repo funding of loans and other financial transactions harder to arrange.

Furthermore, the Fed’s imposition of the Supplementary Leverage Ratio (SLR) requirement has also reduced the supply of repo funding. This policy was announced in late 2012 and became effective in 2013. It includes the quantity of repos (and other items) in the regulatory measure of leverage. In effect, including repo in the SLR means that repo funding is more costly to banks that use it as a source of funding. Allahrakha, Cetina, and Munyan (2016) find that this new requirement significantly increased the cost of repo finance for regulated U.S. institutions.

⁴See the September 2014 FOMC statement, available at www.federalreserve.gov/monetarypolicy/policy-normalization.htm.

The Fed's dual role as regulator and repo counterparty raises important and disturbing questions about a new conflict of interest. As a repo counterparty, the Fed benefits financially from imposing the Supplementary Leverage Ratio, which reduces competitors' abilities to transact in repo. Might the Fed have taken into account its own financial benefits from being able to engage in reverse repo on more favorable terms when setting regulations for its competitors?

When the Fed began contemplating its reverse repo tool (as a means to avoid sales of securities), it was already cognizant that it might want to engage in a large number of such transactions to avoid the political consequences of suffering losses on securities sales and thereby being perceived as contributing to government deficits. I do not claim to know whether the Fed's new SLR rule was motivated in part by a desire to improve its own competitive position in the repo market, but the coincidence in timing between the SLR rule and the Fed's entry into the repo market is disturbing, and there is no question that the Fed suffers a conflict of interest from being both a repo counterparty and a repo regulator.

These conflicts of interest are nothing new. The Fed's regulatory power has long been a lightning rod for politicization, which has often placed the Fed at the center of highly contentious power struggles, often with disastrous consequences for both the economy and the Fed's independence. There are many examples, but the most obvious one has been the Federal Reserve Board's role as the arbiter of bank mergers in the last three decades. The Fed was given that role precisely because it could be counted upon to go along with an ill-conceived government policy, which designed the merger approval process to be a source of rent creation for merging mega banks in the 1990s, and ensured that those rents would be shared between merging banks and urban activist groups, which were given power to influence the merger approval process.

According to Fed officials with whom I have spoken, fears of possible congressional or administrative reprisals against the Fed that might have threatened its monetary policy actions were a major part of the explanation for the Fed's willing participation in this farce. As Stephen Haber and I show in our book, *Fragile by Design: The Political Origins of Banking Crises and Scarce Credit*, Fed bank merger hearings focused on the testimony of activist groups about whether the merging banks were "good citizens," a trait that was measured by the amount of loans and grants the merging banks had

contractually promised to give the activists as the *quid pro quo* for their testimony. Those contractual promises exceeded \$850 billion from 1992 to 2006. The Fed's role in overseeing these unseemly political bargains not only lessened the Fed as an institution, but also helped to precipitate the risky mortgage lending that was at the heart of the recent subprime crisis.

The destabilizing debasement of mortgage standards and prudential bank regulatory standards—which were part and parcel of the political deal the Fed oversaw through its merger powers—profoundly contributed to the financial crisis of 2007–09. If the Fed had not been given the authority to approve mergers and set prudential capital standards, and if merger approval and prudential standards had been based on clear rules enforced by an independent regulatory body, then the subprime crisis might have been avoided, or at least substantially mitigated.

Removing the Fed from its regulatory role would not in any way prevent the Fed from examining banks and pursuing all the related supervisory functions that are necessary to a central bank's lending function. Examination powers and some continued shared supervisory authority should be preserved. But there is no reason for the central bank to determine merger policy, whether banks should be permitted to act as real estate brokers, or other matters unrelated to central banking. And allocating that decisionmaking to the Fed does positive harm by putting the Fed in the line of fire with respect to highly charged political battles, which often results in inferior regulatory decisions and jeopardizes independent monetary policy.

Four reforms would avoid most of the problems that stem from combining the Fed's monetary policy authority with its other authorities and powers. First, the Fed should not hold securities other than U.S. Treasury securities in its portfolio (except briefly in the context of assistance approved under its emergency lending powers).⁵ Second, rather than permit the Fed to set the interest rate paid on reserves, interest on reserves should be fixed at 10 basis points below the federal funds rate. Third, the Fed should be prohibited from competing with other intermediaries in the repo market. Fourth, the 2008 Treasury "blueprint" provided a thoughtful vision of how to

⁵See also Plosser (2017). For a discussion of how to make Fed lender-of-last-resort lending more credibly rule-based, see Calomiris et al. (2017).

reorganize the administration of financial regulation. Avoiding duplication of effort by consolidating regulatory functions (not only in banking, but also by creating a federal charter for insurance companies) is long overdue. This approach also would remove the Fed from the job of writing and enforcing regulations, which would free monetary policy from the conflicts that arise when it is combined with those tasks. The Fed would still participate in examinations and have full access to all information necessary to fulfill its role as a lender of last resort, as envisioned under the Treasury blueprint. At the very least, the Fed should be removed from merger decisions and oversight of highly politically sensitive matters, such as Community Reinvestment Act examinations.

Conclusion

Table 1 summarizes the three sets of reforms proposed in this testimony: internal governance reforms, policy process reforms, and limits on Fed asset holdings and activities. Together these proposed reforms would provide a new approach to governing monetary policy, which would result in better monetary policy decisionmaking than we have witnessed in the troubled first century of the Fed's history (through the channels of influence summarized in Figure 1). A central bank that operates as a more democratic institution, is able to benefit from more diverse thinking, is required to follow transparent and systematic policy actions in pursuit of achievable objectives, is held accountable for its actions, and is freed from myopic political pressures and less conflicted by nonmonetary policy mandates and tools would be much more likely to achieve the proper objectives of monetary policy.

There are practical political considerations that make 2018 an ideal time to push forward on needed reforms. A majority of Republicans have long favored many of the reforms listed here, but they have not been able to gain sufficient support from enough Democrats to enact policy reforms. Over the next two years, all seven Fed governors may be appointed by President Trump and confirmed by a Republican-majority Senate. It seems likely that many Democrats who have opposed Fed reforms in the past now may find it appealing to support measures that would make Fed policymaking more systematic, more receptive to diverse viewpoints, and more immune to political influences.

There is another reform relating to the Fed that should also be implemented, which does not fit into any of the categories discussed above. The Fed's surplus revenues should not be used as an off-budget means of funding the Consumer Financial Protection Bureau, other regulatory activities (including those undertaken by the Fed itself), highway expenditures, or other programs. Those practices undermine honest government budgetary accounting and discipline. If the U.S. government wants to be taken seriously as an instrument of monetary reform, it must also be willing to subject itself to honest accounting.

References

- Allahrakha, M.; Cetina, J.; and Munyan, B. (2016) "Supplementary Leverage Ratio and Repo Supply." Working Paper, Office of Financial Research, U.S. Treasury.
- Calomiris, C. W. (2006) "Alan Greenspan's Legacy: An Early Look: The Regulatory Record of the Greenspan Fed." *American Economic Association Papers and Proceedings* 96 (2): 170–73.
- _____ (2013) "How to Promote Fed Independence: Perspectives from Political Economy and History." Working paper, Columbia University (March).
- _____ (2014) "Diversity Now." *The International Economy* (Winter): 46–49, 83.
- _____ (2017a) *Reforming Financial Regulation*. Manhattan Institute.
- _____ (2017b) "Taming the Two 800 Pound Gorillas in the Room." In R. Bliss and D. Evanoff (eds.) *Public Policy and Financial Economics*, forthcoming.
- _____ (2017c) "The Microeconomic Perils of Monetary Policy Experiments." *Cato Journal* 37 (1): 1–15.
- Calomiris, C. W., and Haber, S. H. (2014) *Fragile By Design: The Political Origins of Banking Crises and Scarce Credit*. Princeton, N. J.: Princeton University Press.
- Calomiris, C. W.; Holtz-Eakin, D.; Hubbard, R. G.; Meltzer, A. H.; and Scott, H. (2017) "Establishing Credible Rules for Fed Emergency Lending." *Journal of Financial Economic Policy* 9 (3): 260–67.
- Eccles, M. (1935) "Testimony of Governor Marriner Eccles of the Federal Reserve Board on Banking." *Hearings before the Senate*

- Committee on Banking and Currency*, 74th Congress, 1st Session (March 4): 179 ff.
- Faust, J., and Henderson, D. W. (2004) “Is Inflation Targeting Best-Practice Monetary Policy?” International Finance Discussion Paper No. 807 (May). Federal Reserve Board.
- Friedman, M. (1948) “A Monetary and Fiscal Program for Economic Stability.” *American Economic Review* 38 (3): 245–64.
- _____ (1959) *A Program for Monetary Stability*. New York: Fordham University Press.
- _____ (1968) “The Role of Monetary Policy.” *American Economic Review* 58 (1): 1–17.
- Kydland, F., and Prescott, E. (1977) “Rules Rather Than Discretion: The Inconsistency of Optimal Plans.” *Journal of Political Economy* 85 (3): 473–91.
- Laeven, L., and Valencia, F. (2013) “Systemic Banking Crises Database.” *IMF Economic Review* 61 (2): 225–70.
- Meltzer, A. H. (2003) *A History of the Federal Reserve, Volume 1*. Chicago: University of Chicago Press.
- _____ (2009a) *A History of the Federal Reserve, Volume 2, Book 1*. Chicago: University of Chicago Press.
- _____ (2009b) *A History of the Federal Reserve, Volume 2, Book 2*. Chicago: University of Chicago Press.
- _____ (2014) “Current Lessons from the Past: How the Fed Repeats Its History.” *Cato Journal* 34 (3): 519–39.
- Orphanides, A. (2003a) “Monetary Policy Evaluation with Noisy Information.” *Journal of Monetary Economics* 50 (3): 605–31.
- _____ (2003b) “The Quest for Prosperity without Inflation.” *Journal of Monetary Economics* 50 (3): 633–63.
- Phelps, E. S. (1968) “Money-Wage Dynamics and Labor-Market Equilibrium.” *Journal of Political Economy* 76 (4): 678–711.
- _____ (1972) *Inflation Policy and Unemployment Theory*. London: Macmillan.
- Plosser, C. I. (2017) “Why the Fed Should Own Only Treasuries.” *Defining Ideas* (June 10). Available at www.hoover.org/research/why-fed-should-only-own-treasuries.
- Taylor, J. B. (1993) “Discretion versus Policy Rules in Practice.” *Carnegie-Rochester Series on Public Policy* 39: 195–214.
- _____ (2016) “Central Bank Models: Lessons from the Past and Ideas for the Future.” Keynote presentation at the workshop,

- “Central Bank Models: The Next Generation,” Bank of Canada (November 17).
- Thornton, D. L., and Wheelock, D. C. (2014) “Making Sense of Dissents: A History of FOMC Dissents.” Federal Reserve Bank of St. Louis *Review* 96 (3): 213–27. Available at research.stlouisfed.org/publications/review/2014/q3/thornton.pdf.
- U.S. Department of the Treasury (2008) “The Department of the Treasury Blueprint for a Modernized Financial Regulatory Structure” (March). Available at www.treasury.gov/press-center/press-releases/Documents/Blueprint.pdf.
- U.S. House of Representatives Committee on Financial Services (2016) “The Financial CHOICE Act: Creating Hope and Opportunity for Investors, Consumers, and Entrepreneurs: A Republican Proposal to Reform the Financial Regulatory System” (June 23). Available at financialservices.house.gov/uploadedfiles/financial_choice_act_comprehensive_outline.pdf.
- Wieland, V.; Afanasyeva, E.; Kuete, M.; and Yoom, J. (2016) “New Methods for Macro-Financial Model Comparison and Policy Analysis.” In J. B. Taylor and H. Uhlig (eds.) *Handbook of Macroeconomics, Volume 2*. Amsterdam: Elsevier.

IMPROVING MONETARY POLICY BY ADOPTING A SIMPLE RULE

Athanasios Orphanides

Monetary theory as well as monetary practice over the past few decades suggest that economic outcomes in our economy are better when monetary policy is systematic and respects the importance of maintaining price stability.¹ Despite broad agreement of the benefits of systematic policy, the Fed continues to set policy on a meeting-by-meeting discretionary basis. This article examines how policy can be improved by replacing discretion with a transparent process of selecting and periodically adapting a simple policy rule.

The Case for a Simple Rule

The Fed's decision to adopt a precise quantitative definition of price stability in January 2012 was an important step in the right direction. With the adoption of an inflation target—2 percent, measured by the PCE index—in its *Statement on Longer-Run Goals and Monetary Policy Strategy* (Federal Reserve Board 2012), the Fed can facilitate well-anchored inflation expectations in line with price stability and can be held accountable over time more easily.

Cato Journal, Vol. 38, No. 1 (Winter 2018). Copyright © Cato Institute. All rights reserved.

Athanasios Orphanides is Professor of the Practice of Global Economics and Management at the MIT Sloan School of Management and a member of the Shadow Open Market Committee (SOMC). This article is a revised version of the author's presentation at the May 5, 2017 meeting of the SOMC (Orphanides 2017).

¹See Taylor and Williams (2011) for a comprehensive review of the literature on policy rules.

However additional progress is required. The Fed's current policy framework places too much emphasis on meeting-by-meeting discretion and is not sufficiently systematic to be in line with best policy practice. This is particularly problematic because of the Fed's so-called dual mandate—to achieve simultaneously maximum employment and price stability.

It is well known that the mandate of the Federal Reserve can create difficulties for the institution when tightening policy is required to keep inflation at bay. The combination of meeting-by-meeting discretion and multiple conflicting goals makes the Fed vulnerable to all the pitfalls that monetary theory and history teach us are associated with the absence of systematic policy. This can be corrected if the Fed adopts and communicates a simple policy rule that it can then use as a guide for setting monetary policy. Adopting an appropriate simple rule would allow the Fed to respond in a countercyclical fashion to economic developments while protecting price stability over time.²

Countering Key Arguments against Rules

Unfortunately, the Fed has not shown the willingness to move in that direction and continues to prefer to operate with meeting-by-meeting discretion. In recent speeches, Fed Chair Yellen and Vice Chair Fischer have presented a case *against* monetary rules (Fischer 2017a, 2017b; Yellen 2017). It is instructive to examine the key arguments presented against rules and provide counterarguments to make progress in this policy debate.

Perhaps the most common argument against monetary rules is that discretion allows greater flexibility to take into account uncertainty. It is certainly important to acknowledge uncertainty. Policymakers face numerous dimensions of uncertainty. Our understanding of how the macro economy works is incomplete. Estimated macroeconomic models are imperfect, and often competing models with quite different policy implications may be equally plausible. Key concepts that would have been very useful for policy, if they could be measured accurately in real-time (e.g., the natural rate of interest), are in fact unknown.

²See Orphanides (2015) for a more detailed exposition.

The presence of uncertainty, however, cannot serve as a valid argument for defending discretionary policy. Indeed, uncertainty raises the potential costs of discretion as it makes it harder to understand how large a policy deviation may be from what would have been the desirable systematic response to a shock.

The reasons why systematic policy is preferable to discretion are no less important under uncertainty. Consider dynamic inconsistency, one of the major policy problems associated with discretionary policy that the adoption of a rule solves. When policy does not follow a rule, households and businesses cannot trust that the policymaker will follow through with any policy that was communicated in the past, even if nothing has changed in the economy. Under discretion, the policymaker has the incentive to deviate from earlier plans and households and businesses must adapt their behavior to protect against these future discretionary decisions. Dynamic inconsistency makes everyone worse off when policy is set under discretion. Dynamic inconsistency is as much a feature of the macroeconomic policy problem under uncertainty about the structure of the economy as it is when, for simplicity, we assume this uncertainty away.

Consider the issues associated with the formation of expectations, most importantly about inflation. A major advantage of monetary rules over discretion is that when the Fed is systematic and follows policy based on a rule, financial market participants, households, and businesses can better understand what the Fed is doing and take that into account in forming expectations. Again, this is the case regardless of whether we assume perfect knowledge about the structure of the economy or acknowledge imperfect knowledge. Systematic policy is even more important when the economy is buffeted by uncertain and potentially destabilizing shocks: when policy follows a well-designed rule, inflation expectations can remain well-anchored, which in turn helps maintain stability both in prices and in the real economy.

Uncertainty is also invoked in another way that is important to address and often used as an excuse to promote discretionary policy. It is *correctly* noted by advocates of discretionary policy that since our knowledge about the structure of the economy is incomplete and our understanding of this structure evolves over time, no simple fixed and immutable monetary rule can possibly be

best at all times. Hence, it is argued, it is best for policy to remain discretionary.

The Taylor rule can be used as an example. Some versions of the Taylor rule state that policy should be set with the implicit assumption that the natural rate of interest is constant. The classic Taylor rule, for example, has embedded in it the assumption that the natural rate of interest equals 2 percent. If policy is set in accordance with this version of the Taylor rule and the wrong assessment of the natural rate of interest, policy would be systematically too easy or too tight, leading to inferior economic outcomes. Thus, it is argued, discretion is better.

This argument, however, is not right. What it suggests is that care is needed in selecting a policy rule that properly accounts for uncertainty, including about the natural rate of interest. Furthermore, it suggests that a simple policy rule should not always be seen as fixed and immutable. Fixed and immutable rules can indeed be problematic if they cannot be adapted as our understanding of the economy evolves.³

As our knowledge improves, we must reevaluate the simple rules-of-thumb embedded in our models and embrace modifications suggested by new analysis. This ought to be the case both for those who argue in favor of formulating policy in a systematic manner and for those who prefer meeting-by-meeting discretion.

For this reason, it is important to describe more precisely a process for selecting a rule that ensures that policy is systematic. To account for our evolving understanding of the economy, the Fed could adopt a framework that relies on a simple policy rule that is subject to periodic reviews and adaptation.

Selecting a Robust Rule

The Fed could select a rule following a rigorous evaluation process that ensures robustness, taking into account all the dimensions of

³Adaptation is required of virtually any rule to avoid systematic errors. This includes interest rate rules and rules regulating the growth rate of the money supply. For example, k -percent money-growth rules require estimates of trend velocity and potential output growth to deliver a 2 percent inflation goal. With no drift in velocity a k -percent rule would suggest 4 percent money growth if potential output growth is believed to be 2 percent but only 3 percent money growth if potential output growth is believed to be 1 percent.

uncertainty that can be incorporated into macroeconometric policy evaluation. The evaluation process should include uncertainty about natural rates, about the structure of the economy, about expectations formation, and so forth.

The Fed could communicate its selected rule, as part of an expanded and more detailed *Statement on Longer-Run Goals and Monetary Policy Strategy*. Adding the Fed's monetary rule to this statement would complete it by actually providing the Fed's policy strategy, which is absent from the current meeting-by-meeting discretionary framework.

It is important to acknowledge that setting policy following a monetary rule is a living process that requires periodic review and adaptation. This would allow the Fed the flexibility to account for and adapt to the evolving understanding of the economy.

The Fed could publish an evaluation of its rule on an annual basis and adapt its rule, if needed. Updates to the Fed's rule could be presented with the annual revision of the *Statement on Longer-Run Goals and Monetary Policy Strategy* that the Fed has published each January since 2012.

Replacing the meeting-by-meeting discretion with a transparent process of selecting and periodically adapting a simple and robust policy rule would ensure that monetary policy is systematic and contributes to social welfare over time while also retaining the flexibility to account for the evolution of the economic environment and of our knowledge. To ensure transparency and accountability, the Fed should communicate its preferred rule with sufficient detail so that an outside observer could track policy using incoming information and data without additional input from the Fed. The detail required would depend on the selected rule. For example, if the rule's implementation required use of unobserved concepts that evolve over time, such as the natural rate of interest, the methodology for arriving at the pertinent estimates should also be specified in advance to make the rule meaningful and avoid discretion.

The framework just described outlines how the Fed could adopt a monetary rule and maintain systematic policy in a manner that addresses the key concerns advanced when a case against monetary rules is argued. It may be noted that congressional legislation could guide the Fed in this direction. For example, the Federal Oversight Reform and Modernization (FORM) Act that was introduced in

Congress and passed by the House in 2015, includes a provision for a Directive Policy Rule which is developed by the FOMC to provide the basis for the Open Market Operations Directive.

Indeed, as the FORM Act implies, the Fed is best placed to select the simple rule that should guide its systematic monetary policy. However, legislation is not necessary for the Fed to adopt a simple rule. The Fed could embrace this improvement on its own, within its current mandate. No change in the Federal Reserve Act is needed for the Fed to include a simple policy rule in its annual *Statement on Longer-Run Goals and Monetary Policy Strategy*. The improvement could be seen as an added step, building on the improvement that started with the adoption of an explicit inflation target in 2012, which did not require a change in the Federal Reserve Act.

Conclusion

I will close by recounting a recent exchange I had with Vice Chair Fischer on this matter. Professor Stanley Fischer, who was one of my teachers at MIT, has been an active participant in the rules versus discretion debate for many decades. In the early 1970s, together with Phillip Cooper, he was among the first to do econometric policy evaluation of monetary rules in competing models such as the FRB-MIT-Penn Model and the St Louis Monetarist model, which were used at the time for policy analysis.⁴ In 1990, he published an influential review of the literature, presenting the rules versus discretion debate that was then “at least 150 years old” (Fischer 1990: 1181). At that time, the research appeared inconclusive and Fischer suggested that a new generation of models needed to be developed. However, in reflecting about what could guide monetary policy in the meantime, he also suggested that “it might

⁴These contributions include Cooper and Fischer (1972a, 1972b, 1974) and Fischer and Cooper (1973). The research program in these papers involved comparative evaluations of active countercyclical monetary rules that were more elaborate than many of the simple Taylor-type rules that have been advocated more recently.

be possible to find simple feedback rules that perform well in a variety of models, and to recommend them as a basis for monetary policy” (p. 1169).

Last February, Vice Chair Fischer delivered a speech arguing in favor of discretion over monetary rules (Fischer 2017a). For the title, he used part of a famous quote by Paul Samuelson highlighting the value of judgement and models for policy analysis. Samuelson said, “I’d rather have Bob Solow than an econometric model, but I’d rather have Bob Solow with an econometric model than without one.” While I fully subscribe both to the use of models and to the value of judgment, I thought that the case presented against rules was incomplete. I wrote an email to the vice chair with the subject line: “I’d rather have Bob Solow with a model and a rule (following a careful evaluation process)” and went on to describe how the Fed could go about selecting a policy rule, relying on the superb research of Fed system staff.

The vice chair responded with a subsequent speech in March (Fischer 2017b). While acknowledging how a careful evaluation process could proceed, he appeared to remain unconvinced and noted, “However, I tend to agree with John Taylor and my Fed colleague John Williams when they write that ‘the search for better and more robust policy rules is never done.’”⁵

Once again, I find myself in agreement with the quote but not with the suggested conclusion. It is indeed true that the search for better rules is never done. It is also true that our knowledge will always remain imperfect. But is this sufficient to justify the Fed’s emphasis on meeting-by-meeting discretion? Perhaps we should acknowledge our imperfect knowledge and promote systematic policy with a robust rule that reflects our current state of knowledge. We could also accommodate potential amendments by embracing a transparent process of periodically adapting the simple and robust policy rule selected to guide monetary policy.⁶

⁵The quote is from the conclusion of Taylor and Williams (2011).

⁶Indeed, such a process would heed the advice: “Prudence . . . suggests that the rule include procedures for its own amendment” (Fischer 1990: 1169).

References

- Cooper, P. J., and Fischer, S. (1972a) “Simulations of Monetary Rules in the FRB-MIT-Penn Model.” *Journal of Money, Credit, and Banking* 4 (2): 384–96.
- _____ (1972b) “Stochastic Simulation of Monetary Rules in Two Macroeconomic Models.” *Journal of the American Statistical Association* 67 (340): 750–60.
- _____ (1974) “Monetary and Fiscal Policy in the Fully Stochastic St. Louis Econometric Model.” *Journal of Money, Credit, and Banking* 6 (1): 1–22.
- Federal Reserve Board (2012) “Federal Reserve Issues FOMC Statement of Longer-Run Goals and Policy Strategy.” Press Release (January 25).
- Fischer, S. (1990) “Rules versus Discretion in Monetary Policy.” In B. M. Friedman and F. Hahn (eds.) *Handbook of Monetary Economics*, Vol. 2, 1155–84. Amsterdam: North-Holland.
- _____ (2017a) “I’d Rather Have Bob Solow Than an Econometric Model, But” Remarks at the Warwick Economics Summit, Coventry, U.K. (February 11).
- _____ (2017b) “Monetary Policy: By Rule, By Committee, or By Both?” Remarks at the 2017 U.S. Monetary Policy Forum, New York (March 3).
- Fischer, S., and Cooper, P. J. (1973) “Stabilization Policy and Lags.” *Journal of Political Economy* 81 (4): 847–77.
- Orphanides, A. (2015) “Fear of Liftoff: Uncertainty, Rules, and Discretion in Monetary Policy Normalization.” Federal Reserve Bank of St Louis *Review* 97 (3): 173–96.
- _____ (2017) “The Case Against the Case Against Monetary Rules.” Position paper for the May 7, 2017 meeting of the Shadow Open Market Committee.
- Taylor, J. B., and Williams, J. C. (2011) “Simple and Robust Rules for Monetary Policy.” In B. M. Friedman and M. Woodford (eds.) *Handbook of Monetary Economics*, Vol. 3B, 829–60. Amsterdam: North-Holland.
- Yellen, J. (2017) “The Economic Outlook and the Conduct of Monetary Policy.” Remarks at the Stanford Institute for Economic Policy Research, Stanford (January 19).

AGAINST HELICOPTER MONEY

Kevin Dowd

Back in 1969, Milton Friedman proposed an interesting thought experiment that has since become famous:

Let us suppose now that one day a helicopter flies over this community and drops an additional \$1,000 in bills from the sky, which is, of course, hastily collected by members of the community. Let us suppose further that everyone is convinced that this is a unique event which will never be repeated [Friedman 1969].

Friedman did not intend his suggestion as a serious policy proposal. Instead, he intended it as a classroom device to illustrate the consequences of changes in the stock of base money. The idea then stayed in the classroom for many years, virtually unknown to all except academic monetary economists.

In the late 1990s, it began to be reinvented as a serious policy proposal. People first began to think it might be a useful instrument to combat deflation in Japan. The idea then hit the headlines in 2002 when then-Fed governor Ben Bernanke suggested that it might be used to combat possible deflation in the United States too:

The U.S. government has a technology, called a printing press (or, today, its electronic equivalent), that allows it to produce as many U.S. dollars as it wishes at essentially no cost.

Cato Journal, Vol. 38, No. 1 (Winter 2018). Copyright © Cato Institute. All rights reserved.

Kevin Dowd is Professor of Finance and Economics at Durham University. He thanks Frank Browne, Dave Cronin, Jim Dorn, Anthony Evans, Martin Hutchinson, and Basil Zafiriou for helpful comments. The usual caveat applies.

By increasing the number of U.S. dollars in circulation, or even by credibly threatening to do so, the U.S. government can also reduce the value of a dollar in terms of goods and services, which is equivalent to raising the prices in dollars of those goods and services. We conclude that, under a paper-money system, a determined government can always generate higher spending and hence positive inflation [Bernanke 2002].

He went on to make a passing reference to Friedman's helicopter money idea: "A money-financed tax cut is essentially equivalent to Milton Friedman's 'helicopter drop' of money," he said. That passing reference was enough to saddle him with a nickname—"Helicopter Ben"—which he has been stuck with ever since.¹

The attraction of the idea to its proponents is the promise of being able to deliver monetary stimulus in circumstances in which interest rates are low or zero and in which the traditional tools of monetary stimulus might be ineffective or infeasible.

Then came the financial crisis, and one unconventional stimulus policy after another—quantitative easing (QE), zero interest rate policy, and, in some countries, negative interest rate policy—failed to deliver the desired results. As a consequence, helicopter money is enjoying a new revival.

For some of its advocates, the core argument is that we have tried everything else to boost the economy and helicopter money is or might be the only instrument of stimulus left in the central bank's toolbox (see, e.g., Perotti 2014; Turner 2015). Others argue that helicopter money should be considered if fiscal stimulus is politically inexpedient (e.g., Brittan 2012), as a tool to counter income inequality (Muellbauer 2014) or a blocked transmission mechanism (see Durden 2015), or to pull the economy out of deflation (e.g., Bernanke 2002; Selgin 2016), secular stagnation (e.g., Buiter, Rahbari, and Seydl 2015), or a liquidity trap (e.g., Caballero 2010). Some advocates of helicopter money combine a number of these arguments. A prominent example is financial journalist Martin Wolf,

¹In his memoirs, Bernanke (2015: 64) comes close to suggesting that he might have made a mistake in using the helicopter label: "Dave Skidmore, the media relations officer . . . had advised me to delete the helicopter-drop metaphor. . . . 'It's just not the sort of thing a central banker says,' he told me. I replied, 'Everybody knows Milton Friedman said it.' As it turned out, many Wall Street bond traders . . . apparently [did] not."

who suggests that in the absence of more fiscal stimulus, which would be politically problematic, “central banks are the only players” left:

policymakers must prepare for a new “new normal” in which policy becomes more uncomfortable, more unconventional, or both. Can the world escape from the chronic demand weakness? Absolutely, yes. Will it? That demands greater boldness. When one has exhausted the just about possible, what remains, however improbable, must be the answer [Wolf 2016].

The “what remains” he had in mind is helicopter money. Yet the idea appalls critics. Helicopter money breaks the ultimate monetary taboo against the wanton overissue of base money. It harnesses central bankers’ most primitive power, their unique ability to create base money at will and at negligible cost, and proposes to create as much base money as it takes to achieve the desired boost to spending. It is the monetary weapon of last resort—a weapon so potent that central bankers have traditionally feared to deploy it lest it destroy the currency and potentially much else besides. To quote the British journalist Ambrose Evans-Pritchard (2013): “A great many readers in Britain and the U.S. will be horrified that this helicopter debate is taking place at all, as if the QE virus is mutating into ever more deadly strains.”

I begin this article by explaining how helicopter money should be accounted for on the central bank’s balance sheet. I then explain the basics of helicopter money and compare it to some alternatives. After that, I give some examples of helicopter money, explain how it impacts the central bank’s balance sheet, and point out that—appearances notwithstanding—helicopter money is not a “free lunch.” Finally, I discuss the (significant) problems entailed by helicopter money, concluding that traditional monetary economists were right to warn against it.

Accounting for Helicopter Money

We should first consider how to account for helicopter money on the central bank’s balance sheet. In conducting helicopter money operations, the central bank issues additional base money that it gives away: it increases its liabilities without receiving any valuable asset in return. At the same time, its balance sheet must always balance. So how do we reconcile these statements? The most natural way to do

so is to imagine that in conducting helicopter money operations the central bank receives a hypothetical asset that has a book value equal to the amount of newly issued base money but that has a market value of zero.

An asset that meets these requirements is a perpetual bond with a zero coupon. Perpetual means that the bond will never mature, and the zero coupon implies that there will be no regular coupon payments—that is, the bond promises no payments at all. The book value of the bond must be equal to the amount of newly issued base money in order for the central bank balance sheet to balance, but its market value will be zero because no one would pay anything for a bond that offers no future payments.

We can then think of helicopter money as an operation in which the central bank issues additional base money in return for which it receives a zero-coupon perpetual bond, and it helps to think of this bond as being issued by or on behalf of the recipients of the central bank's newly issued base money. Of course, such a bond is a fiction, but it is an illuminating one as we shall see.

The Basics of Helicopter Money

Helicopter money is often described as the central bank “printing” money and giving it away. This description captures the essence of the idea but is a little misleading. Most proposals would involve the central bank issuing and giving away not physical cash, but a digital equivalent. The central bank might issue every citizen with the digital or credit equivalent of \$500 or whatever, which might be delivered directly to individuals' bank accounts. The key point is that the newly issued base money would be issued directly to the beneficiaries.

Those beneficiaries might be members of the public, but under some helicopter money proposals the “helicopter drops” might be targeted toward approved projects—in infrastructure, for example.

From the beneficiaries' perspective, the payments are free and therefore naturally welcome. However, decisions need to be made over who the beneficiaries should be and how much they should be paid. These considerations imply that helicopter money has a redistributive element, but redistribution is traditionally regarded as falling under the domain of *fiscal* policy. I will return to this issue below.

In its typical forms, helicopter money works to increase aggregate demand: people pick up the dollar bills and spend some of

them. We should not assume, however, that all the helicoptered money would be spent: according to standard models of consumption and saving, individuals' marginal propensities to spend will typically be less than 100 percent, and will also depend on factors such as their expectations of future income and inflation. The likely impact would then be some increase in aggregate demand, which would produce some increase in output and some increase in prices. In the longer term, the output stimulus is likely to wear off and the long-run outcome would be a higher price level. Repeated helicopter money operations would then take us into the familiar world of the expectations-augmented Phillips curve set out by Friedman (1968) in his presidential address to the American Economic Association.

The success or otherwise of helicopter money operations depends, therefore, on their purpose. If their purpose is to stimulate output, they would likely only have some short-term success. But if their purpose is to increase aggregate demand or prices (or, if applied in repeated doses, inflation), then helicopter money could be a potent monetary policy instrument—for better or for worse. To illustrate: Willem Buitier (2014: 2) demonstrates that under very general conditions, including a permanent liquidity trap and even Ricardian Equivalence,

there always exists a combined monetary and fiscal policy action that boosts private demand—in principle without limit. Deflation, inflation below target, “lowflation,” “subflation” and the deficient demand-driven version of secular stagnation are therefore unnecessary. They are policy choices. This effectiveness result holds when the economy is away from the zero lower bound (ZLB), at the ZLB for a limited time period or at the ZLB forever.

This suggests that the policy question is not whether helicopter money *would* increase demand or prices, but instead whether we should ever *want* to use it to achieve those outcomes.

Helicopter Money Compared to Alternatives

We can gain further insights into helicopter money by comparing it to several familiar alternatives—QE and debt monetization—as well as to a less familiar alternative, gold stock revaluation.

Helicopter Money Compared to Quantitative Easing

QE is the policy of issuing base money to finance large-scale purchases of financial assets. QE is meant to support the banking system and to reduce interest rates. Payments are typically to financial institutions and not directly to members of the public. By contrast, helicopter money typically involves payments to the public or, in some cases, payments to finance specific projects. It is not usually (if ever) proposed to support the banking system or to reduce interest rates.

Helicopter Money Compared to Debt Monetization

Debt monetization occurs when the Treasury issues a bond that it sells to the central bank, which in turn pays for the bond with newly issued base money. The bond is then used by the Treasury to make payments in pursuance of government fiscal policy objectives.

In both helicopter money and debt monetization operations, the central bank issues new base money and acquires a bond, at least hypothetically. However, these operations differ in three important respects.

First, under debt monetization the bond must promise at least one future repayment and have a positive market value; in contrast, the hypothetical “bond” that the central bank obtains under helicopter money provides for no future payments whatsoever and has a market value of zero.

A second difference relates to who makes the decisions about who the beneficiaries should be and how much they should be paid. Under debt monetization, these decisions are always made by the government or by its fiscal agents in the pursuit of government fiscal policy objectives; under many (though not all) versions of helicopter money, such decisions are made by the central bank.

Finally, debt monetization and helicopter money face different political and legal constraints. There are political costs to advocating expansionary policies such as large-scale debt monetization, which can be criticized as fiscally irresponsible. There are also constraints against fiscal expansion (such as the United States’ federal debt ceiling) that might scupper debt monetization, but can be circumvented by using helicopter money. That ability of helicopter money to circumvent political and legal constraints is a key argument used by both proponents and opponents of helicopter money. The issue here is whether one sees such constraints as

inconvenient barriers to be overcome or as important bulwarks to be protected.

Helicopter Money Compared to Gold Stock Revaluation

The Fed currently values its gold stock—which was equal to about 13.5 million fine Troy ounces of gold as of September 16, 2016—at \$42.22 per ounce. If it wished to, the Fed could revalue these at a price closer to the current market price, which was about \$1,312 an ounce at that same date. As Jerry Jordan (2016) points out, the Fed carried out such operations twice more than four decades ago. At the end of 1971, the “official” price of the U.S. stock of gold in Fort Knox was raised from \$35 an ounce to \$38 an ounce, and two years later it was raised again to \$42.22 an ounce, a value that still stands. On both occasions, the asset value of the “gold certificates” held by the Fed—its claims to the gold held in Fort Knox—was increased, so its liabilities had to rise by the same amount. The U.S. Treasury “General Account” at the Fed was then credited with the amount of the increase in the value of U.S. gold holdings, so that the Treasury could spend that amount without having to collect taxes or sell bonds. Earlier still, the Gold Reserve Act of 1934 had raised the official price of gold from \$20.67 to \$35 per Troy ounce and produced a similar “windfall gain” for the government.

Such operations meet Jordan’s definition of helicopter money as an increase in central bank liabilities without an open market purchase of securities by the central bank. However, in this gold revaluation operation there is only one beneficiary—the Treasury—whereas under my definition of helicopter money the beneficiaries are the public or some set of private-sector parties.

There is also another difference: under helicopter money by my definition, the Fed’s assets do not increase in value, whereas, under an upward gold revaluation of the sort Jordan discusses, the Fed’s asset holdings increase in value.

Examples of Helicopter Money

Helicopter money can take many different forms. Leaving aside Friedman’s famous helicopter drop, obvious examples include direct cash handouts to the public, sending people checks in the mail, or making direct payments to people’s bank accounts. A variation on this theme would be to issue tax rebates: the Fed might issue

vouchers to taxpayers that they could redeem with the IRS and that the IRS could redeem at the Fed for newly issued base money.

An earlier example comes from Keynes's *General Theory*, which suggested that the central bank should print money that would be used to finance housing projects:

If the Treasury were to fill old bottles with banknotes, bury them at suitable depths in disused coalmines which are then filled up to the surface with town rubbish, and leave it to private enterprise on well-tryed principles of laissez-faire to dig the notes up again (the right to do so being obtained, of course, by tendering for leases of the note-bearing territory), there need be no more unemployment and, with the help of the repercussions, the real income of the community, and its capital wealth also, would probably become a good deal greater than it actually is. It would, indeed, be more sensible to build houses and the like; but if there are political and practical difficulties in the way of this, the above would be better than nothing [Keynes 1936: 129].

Keynes's proposal illustrates that helicopter money can be used to finance "worthwhile" projects, however those might be defined. More precisely, it shows that helicopter money can be used to finance projects that *somebody important* regards as worth financing, but that presumably would not otherwise obtain finance.

A more recent example of a helicopter money proposal to finance infrastructure projects is the pan-European "QE for the People" campaign, which proposes "that the money currently being created by the European Central Bank should be given directly to eurozone citizens and/or spent on much needed public investment such as green infrastructure, affordable housing—or in any other way which would contribute to the genuine development of the real economy."² Despite the presence of the term "QE" in the title, this proposal is classic helicopter money, and not QE in the usual sense of a program of large-scale asset purchases from the banks.

Such proposals should be distinguished from superficially similar-sounding proposals such as "Green QE" or "People's QE." The former term is usually used to refer to bond-financed infrastructure projects, especially those advocated by environmentalist groups

²See www.qe4people.eu/about. See also Muellbauer (2014).

(e.g., Murphy and Hines 2010). The latter refers to U.K. Labour Party leader Jeremy Corbyn’s proposal for a bond-financed U.K. National Investment Bank that would invest in politically favored projects. If we categorize these proposals using the QE vs. debt monetization vs. helicopter money trichotomy set out earlier, proposals for “Green QE” and “People’s QE” are essentially proposals to use debt monetization to finance approved infrastructure projects.³

Helicopter Money and the Fed’s Balance Sheet

Helicopter money has important implications for the Fed’s balance sheet. As of September 16, 2016, the Fed’s equity capital was \$40.2 billion and its total liabilities (which are also equal to its total assets) were \$4,441.5 billion—or almost \$4.5 trillion. Its capital to assets ratio was therefore 0.9 percent and its leverage was $4,441.5/40.2 = 110.5$. The former would be regarded as extremely low and the latter extremely high for any financial institution other than a central bank.⁴

Suppose the Fed engaged in a helicopter money operation in which it issued \$1 billion. Its liabilities—the supply of base money—would increase by \$1 billion, but the (market) value of its assets would remain the same. The (market) value of the Fed’s capital, the difference between the value of its liabilities and the (market) value of its assets, would then fall by \$1 billion.⁵

³Those who advocate using helicopter money to finance pet projects generally fail to address the question of why their preferred “socially worthwhile” projects are not already receiving funding. That is, they fail to demonstrate how their proposals would improve investment decisionmaking, they fail to address the knowledge and incentive problems faced by government-funded bureaucrats making investment decisions, and so on. For more on these problems, see Lacalle (2017).

⁴A caveat, however: Recall that the official price of gold is still only \$42.22 an ounce, and yet the market price at the same date was about \$1,312 an ounce. This price difference implies that the Fed’s gold certificates were undervalued relative to market valuations. Since each ounce of gold was undervalued by almost \$1,270 and the Fed had certificates to claim 13,452,810.53 ounces, the Fed’s gold certificates were undervalued by almost \$17.1 billion. If we apply this revaluation, then the Fed’s capital would have been about \$57.3 billion, its assets would have been \$4,441.5 billion + \$17.1 billion = \$4,458.6 billion, and its leverage would have been 77.8, a level which is still high by the standards of other financial institutions.

⁵The book value of the Fed’s assets would presumably increase by \$1 billion to balance the Fed’s books, but the book value is essentially irrelevant and in the text below I use the term “capital” to refer to market-value capital only.

Now suppose that the Fed engaged in a helicopter money operation on a scale approaching \$40.2 billion. Its capital and therefore its capital to assets ratio would then approach zero, while its leverage would soar toward infinity. If the helicopter money operation exceeded \$40.2 billion, then the Fed's capital would become negative and the Fed would be technically insolvent.

However, \$40.2 billion is a relatively small amount in central banking terms. To have a major impact, any Fed helicopter money operation would need to be on a much bigger scale. Given the scale of its QE operations—these amounted to more than \$3.7 trillion—we could envisage helicopter money also on a scale of trillions. For example, with a helicopter money operation of \$1 trillion, say, the Fed's capital would fall to about *minus* \$960 billion. The question then is whether the impact of helicopter money on a central bank's capital or leverage really matters.

Now, for any commercial bank—or indeed, any firm other than the central bank itself—the size of its capital relative to its assets (or the inverse, the ratio of assets to capital, i.e. leverage) matters greatly. Suppose a commercial bank embarked on a policy of repeatedly purchasing assets that then lost all their value. Each additional loss is a deduction from its assets and (therefore) an equal deduction from its capital as well. Since capital is less than assets, the bank's capital to assets ratio would keep falling and the bank's leverage would keep rising. A point would eventually come when the bank would lose the confidence of its creditors, who would then run it out of business.

However, under a monetary regime in which a central bank issues an irredeemable fiat currency, there is a sense in which central bank leverage does not matter. In principle, a central bank can issue as much helicopter money as it likes, take the losses on its balance sheet, and allow its capital to go far into negative territory. What matters here is that when a central bank that issues irredeemable currency takes ever-mounting losses, there is no mechanism equivalent to a bank run to force it to change its policy, let alone run it out of business, because its currency is irredeemable and it can issue more at will. From this perspective, a central bank's capital or leverage numbers are operationally irrelevant. Unlike any other financial institution, a central bank can have arbitrarily high leverage and even negative capital without any impairment to its normal operational functions.

Does this mean that central bank capital and leverage *never* matter? Not at all. The Fed's having negative capital would likely have

some adverse impact on public trust in the Fed, and one could imagine awkward questions from critical members of Congress.⁶ It would further erode international confidence in the dollar too. Consider also two hypothetical examples:

Imagine that the Fed had been stripped of its currency issue privileges and the United States had returned to a gold standard. The Fed would now issue a redeemable currency and face competition from other banks that would be free to issue their own gold-backed currency. Under these circumstances, the Fed would be no different from other banks and would need to maintain a competitive capital position to survive in the long term. Its current high leverage would then be a substantive concern.

As a second example, suppose the Fed issued trillions in helicopter money but later decided to peg the dollar to some other currency or to gold, and to convert the Fed into a currency board. The currency board arrangement would require the Fed to maintain a reserve ratio of at least 100 percent against its base money issue. However, the Fed would now be trillions short in assets, and the only feasible way in which that shortfall could be made good would be for the government to recapitalize the Fed at public expense.

I don't wish to argue the merits, demerits, or likelihood of such reforms here. My point is merely that while central bank capital and leverage might not affect the way that a central bank operates under current institutional arrangements, they have major implications under alternative institutional arrangements.

Is Helicopter Money “Free”?

It is often suggested that helicopter money is somehow “free.” Now, helicopter money is obviously “free” as far as its recipients

⁶The Fed would appear to have anticipated this problem. Writing on April 8, 2011, John P. Hussman (2011) stated: “To avoid the potentially untidy embarrassment of being insolvent on paper, the Fed quietly made an accounting change several weeks ago that will allow any losses to be reported as a new line item—a ‘negative liability’ to the Treasury—rather than being deducted from its capital. Now, technically, a negative liability to the Treasury would mean that the Treasury owes the Fed money, which would be, well, a fraudulent claim, and certainly not a budget item approved by Congress, but we’ve established in recent quarters that nobody cares about misleading balance sheets, Constitutional prerogative, or the rule of law as long as speculators can get a rally going, so I’ll leave it at that.”

are concerned: they get something for nothing. However, to say that something is “free” is to suggest that a valuable asset can be conjured up out of nothing—that is, it has no opportunity cost. But in the case of helicopter money, the opportunity cost is lost seigniorage—the forgone profits from the issue of base money.

To appreciate this point, consider that when a central bank engages in a helicopter money operation it acquires a notional bond “asset”—our earlier zero-coupon perpetual bond—that promises no future payments. However, the central bank could instead have issued the same amount of additional base money to acquire a bond that did promise future payments. For example, the central bank could have acquired a conventional bond issued by a private-sector institution. The payments from this bond are forgone seigniorage profits from helicopter money.

One might object that these forgone seigniorage profits do not matter to the central bank because it hands over its profits to the Treasury. This claim is true but merely confirms my argument. Those forgone seigniorage profits would still matter because they would be lost to the Treasury. Therefore helicopter money has an opportunity cost and is not “free.”⁷

That helicopter money is not “free” can also be seen by comparing the impact of debt monetization and helicopter money on the consolidated government balance sheet, which considers the Fed as part of the government. Consider two cases:

Case 1 (debt monetization) is where the government runs a fiscal deficit that it finances by selling a perpetual coupon-paying bond, or Consols bond, to the central bank, which issues base money to purchase it.⁸

⁷This argument about helicopter money not being free holds true even if interest rates are zero or negative. Suppose that we have zero interest rates and these lead to bonds having zero coupons. In this case, the forgone opportunity is for the central bank to have acquired a zero-coupon bond that repays its face value on its maturity date. That terminal payment is forgone seigniorage. In the case of bonds with negative coupons, the same terminal payment will still be forgone seigniorage, albeit now offset somewhat by the negative coupon payments. However, it would still be the case that the present value of the terminal payment would outweigh the (negative) present value of the coupon payments—otherwise the bond would have no positive price—and so those forgone payments would still have a positive present value. In other words, there is still forgone seigniorage.

⁸These bonds were used by British governments to finance their 18th century and Napoleonic wars. For more on this subject, see Hutchinson and Dowd (2017).

Case 2 (helicopter money) is where the government runs a fiscal deficit that is financed by the central bank issuing base money. This operation involves the central bank's making a credit transfer to the Treasury's central bank account. This case is equivalent to the government's running a deficit that it finances by "selling" the central bank a zero-coupon perpetual bond at its market value of zero.

The only difference is that in Case 1 the government makes subsequent coupon payments to the central bank, whereas in Case 2 it does not. In Case 1, the government makes the coupon payments to the central bank and the central bank remits its seigniorage profits to the government. In Case 2, the government makes no coupon payments to the central bank and the central bank remits its profits to the government, but those profits are less by exactly the amount of the coupon payments in Case 1. Consequently, both operations have an identical impact on the consolidated government balance sheet.

No one would argue that Case 1 involved "free" money or a zero opportunity cost, but since the two cases are functionally indistinguishable, one cannot argue that the former has an opportunity cost but the latter does not.

To repeat: helicopter money is not a free lunch because it has a tangible opportunity cost—the cost of the seigniorage forgone.

John Kay (2016) offers a complementary explanation for why helicopter money is not a free lunch:

For every credit there is a corresponding debit; for every financial asset there is a corresponding liability. The double entry principle is [as] immutable as the first law of thermodynamics, to which it bears a certain resemblance.

You cannot create financial assets out of nothing, just as you cannot create energy out of nothing. There is no financial free lunch, just as there is no perpetual motion. The laws of thermodynamics render perpetual motion machines impossible, though it may be hard to identify the defect in any particular device. And similarly, schemes which purport to fund government spending or tax cuts at no cost to present or future taxpayers are necessarily fatally flawed, even if it requires a little effort to work out the flaw.

To cut to the chase, if we think of the central bank as consolidated within the government, the issue of base money is not a "free" good for the government, because base money can always be used to discharge one's obligations *to* the government—for example, to pay taxes.

However, an objection should be considered: namely, that helicopter money is “free” because it produces some net benefit. Bossone (2016) offers an example of this reasoning, arguing that helicopter money “is a free lunch in the simple sense that, if it works and succeeds in closing the output gap, people won’t have to repay it through higher taxes or undesired (above optimal) inflation.”

I do not buy this argument. The problem with it is apparent from the word “if”: it presupposes that helicopter money *will* have the effects hoped for by its proponents, when in fact no such certainty exists. As a counterexample, consider John Law’s Mississippi scheme, which produced a short-term boom but had ruinous long-term effects; it might have looked like a free lunch while it was in play, but with the benefit of hindsight we see things very differently (see, e.g., Hutchinson 2017). Put more broadly, Bossone’s argument could be used to justify any macroeconomic policy as providing a free lunch, so long as the advocates of that policy merely assert that it will produce some net benefit. Every economic policy ever proposed would then become a free lunch, however destructive it later turned out to be.

More specifically, this argument confuses a genuine free lunch with a questionable net benefit, and these are not the same. The one is tangible and quantifiable; the other involves a speculative projection and may not even exist. To confuse the two involves a category mistake, a logical fallacy.

Problems with Helicopter Money

Helicopter money and debt monetization policies are very similar: they both involve increases in base money, and they have the same impact on the consolidated government balance sheet. Yet there are also differences that are important to note.

Helicopter Money Is Dominated by Debt Monetization

Comparing the two, advocates of helicopter money sometimes argue that it is politically easier to sell when there is a Republican administration, an argument that questionably assumes that a Republican administration is more averse to fiscal stimulus, or that helicopter money is to be preferred because it gets around the constraint of a federal debt ceiling. Such arguments are merely ones of political convenience and are not based on any underlying principle of political economy. Those who make such arguments also tend

to emphasize that helicopter money is “just” a helpful “technical” policy instrument that can be finely calibrated by expert central bankers and their advisors. They play down fears of uncontrolled “money printing” as atavistic and exaggerated.

There are a number of counterconsiderations, however. Such arguments for helicopter money depend on an underlying presumption that central bankers can be trusted to know what they are doing; yet central bankers’ performance over the last decade must surely call their credibility into question. One can also argue that, of the two, debt monetization is to be preferred for reasons of transparency and government fiscal accountability, and that there are no good (e.g., economic or institutional) arguments to prefer helicopter money to debt monetization, as opposed to mere arguments of expediency. Above all, such arguments for helicopter money ignore the reasons fiscal and monetary institutions were designed to have inbuilt constraints in the first place: the terms “division of powers” and “rule of law” come to mind. These helicopter money arguments ignore the reasons why monetary and fiscal functions were intended to be separated, why Congress legislated a debt ceiling, and so forth. Such institutional arrangements were designed to protect the checks and balances in the system. Those who wish to overturn them seek to ride roughshod over the economic constitution, as it were, as if it served no purpose worth defending.

To say that debt monetization is to be preferred to helicopter money is by no means to endorse debt monetization. Large-scale debt monetization has a bad reputation itself, and for good reason: it conjures up images of the Weimar Republic and other slides into hyperinflation, where the failure to respect institutional constraints led to disaster.

Let’s return to the difference between the two policies. The difference is that debt monetization involves an explicit increase in the federal government’s indebtedness, whereas under helicopter money that same increased indebtedness is written off by the Fed. But that difference also has a name: the trillion dollar coin, the proposal that the Mint issue a trillion dollar coin and hand it over to the Fed, which would credit the Treasury’s account with a trillion dollars to spend.⁹ The most prominent advocate of the trillion dollar

⁹The trillion dollar coin proposal seeks to exploit a legal loophole that allows the Treasury to mint platinum coins in any denominations it chooses.

coin is Paul Krugman, who supports it as a means to get around the federal debt ceiling and assures us that it would do “no economic harm at all” (Krugman 2013). Note the underlying assumptions: (1) that the debt ceiling serves no useful purpose; and (2) that allowing greater government deficit spending does absolutely no harm. If one agreed with those assumptions, then Krugman’s argument would be hard to fault. I am inclined myself to agree with the Bank of America analyst who described the proposal as “the latest bad idea,” a “trillion dollar tooth fairy” straight from “the land of fiscal make believe” (quoted in Durden 2013). Krugman is however to be commended for openly acknowledging that the argument for the trillion dollar coin is purely one of political convenience.

Fiscal Problems of Helicopter Money

I have already noted that, in conducting helicopter money operations, the central bank would be making redistribution decisions that are fiscal in nature and therefore more properly belong under the domain of the government. David Stockman (2016) goes further and suggests that helicopter money is

a central bank power grab like no other because it insinuates our unelected central bankers into the very heart of the fiscal process.

Needless to say, the framers delegated the powers of the purse—spending, taxing and borrowing—to the elected branch of government, and not because they were wild-eyed idealists smitten by a naïve faith in the prudence of the demos.

To the contrary, they did so because the decision to spend, tax and borrow is the very essence of state power. There is no possibility of democracy—for better or worse—if these fundamental powers are removed from popular control.

Helicopter money thus raises issues of profound constitutional importance. Bundesbank President Jens Weidmann (2016) expresses similar concerns but also warns that helicopter money could rebound on the central bank itself:

The question of whether and how money is given away to the general public is a highly political one that would need to be addressed by governments and parliaments. Central banks don’t have a mandate to do so, not least because it would

mean redistributing assets on a huge scale. It would be nothing short of unreservedly commingling monetary and fiscal policy, a step which would be incompatible with the notion of central bank independence.

Having the central bank make fiscal policy decisions is to disturb a delicate balance and sow the seeds of conflict between the fiscal and monetary arms of government. A likely consequence is a reaction by Congress that could eventually lead to greater government control over the Fed and the potential end of its independence.

Most proponents of helicopter money acknowledge this monetary/fiscal overlap issue but tend to see it merely as a “coordination” problem.¹⁰ However, such acknowledgements misdescribe the underlying problem as a technical one and in any case do not offer any practical solution to it. The problem is not so much “coordination” as power, because there are underlying issues of control at stake in a four-way turf war between the administration, Congress, the Treasury, and the Fed.

One “technical” solution to this problem, considered merely as a coordination issue, is proposed by Bernanke (2016) himself:

Ask Congress to create, by statute, a special Treasury account at the Fed, and to give the Fed (specifically, the Federal Open Market Committee) the sole authority to “fill” the account, perhaps up to some prespecified limit. At almost all times, the account would be empty; the Fed would use its authority to add funds to the account only when the FOMC assessed that an MFFP [Money-Financed Fiscal Program, i.e., helicopter money] of specified size was needed to achieve the Fed’s employment and inflation goals.

¹⁰There are also other problems with this fiscal/monetary “coordination” issue. The first are legal impediments to central banks financing their governments. Brehon and Winkler (2016) provide an overview of these barriers across the major western economies and conclude that these barriers are not as difficult to surmount as they might appear to be on paper—hardly a reassuring conclusion for those concerned about the protections against abuse that those legal barriers were meant to provide. Another problem is that the Bernanke solution of a designated Treasury account at the central bank is not always easily transferable to other countries. A case in point is Europe: if the ECB were to attempt to implement such a solution, it would soon provoke a major argument over which government or governments should have access to it—the EU Commission, the national governments of the EU, or the eurozone member states.

Should the Fed act, under this proposal, the next step would be for the Congress and the Administration—through the usual, but possibly expedited, legislative process—to determine how to spend the funds (for example, on a tax rebate or on public works). . . . Importantly, the Congress and Administration would have the option to leave the funds unspent. If the funds were not used within a specified time, the Fed would be empowered to withdraw them.

Bernanke's proposal would work as a technical means of implementing helicopter money, although whether it would deliver the results he desires is a different matter.

The idea is that the Fed sets out the size and terms of the operation and the grateful Treasury decides how the helicopter money is to be spent. Yet this proposal is naïve in a number of respects. First, it envisages Congress granting the Fed sole authority to fill the account; however, one can imagine debates in Congress in which it is proposed that Congress should have at least some say and perhaps the ultimate say in this matter. We would then have the potential for a major clash of control between Congress and the Fed that would likely end in the Fed losing power to Congress.

Second, it envisages the possibility of Congress and the administration leaving the funds unspent. But under what circumstances can we plausibly imagine the cash-strapped government *not* making full use of “free” helicopter money provided by the Fed? Let's face it: whatever “free” money the Fed creates for Congress, Congress will spend it all and then come back for more.

Instead, the Bernanke proposal provides Congress with what Stockman describes as “a purportedly scientific monetary cover story” that amounts to a license to spend without having to go through the hassle and unpopularity of voting for taxes to finance it. As he puts it:

The peoples' elected representatives would relish this “expert” cover for ever bigger deficits and the opportunity to wallow in the pork barrel allocation of the targeted tax cuts and spending increases. There is not a hard core New Dealer turning in his grave who could have imagined a better scheme for priming the pump. . . . What makes helicopter money so positively insidious is that it relieves elected politicians entirely from their vestigial fears of the public debt and from accountability for the burdens it imposes on future generations [Stockman 2016].

Under the Constitution, Congress—and only Congress—is authorized to make decisions about fiscal policy. Government spending is constrained by the need for Congress to find some means to finance its spending decisions. It is politically costly for Congress to raise taxes: legislators get flak and they can lose their seats. And so it should be: to finance any government spending, resources have to be extracted—directly via taxation or indirectly via helicopter money—from the public using the coercive powers of the state. It is therefore important that that power not be overused.

But under helicopter money legislators would face the prospect of being able to finance their favorite projects at no political cost to themselves. Since that finance does have a cost, helicopter money would further increase the incentive to generate excessive government spending. Indeed, since helicopter money finance appears to be free and would in fact be free to the legislators who make the key decisions, there would potentially be no limit to the demand for such finance: the supply of (supposedly) worthwhile projects to be financed at (supposedly) zero cost would be infinite.

You cannot have rational spending decisions in a system in which finance appears to be free and legislators are spending money like it grows on trees: rational fiscal policy is impossible when legislators operate under the illusion of a free lunch.

Put it this way: throughout history, governments have faced the problem that taxes were unpopular—poll taxes provoked the Peasants’ Revolt, tea taxes provoked the Boston Tea Party, and so on; then someone comes along who offers them a tax-free solution that enables them to spend as they wish, a fiscal Holy Grail. Do we really expect those in Congress to pass that up because they prefer the unpopularity of taxation? Well, even if they did, they surely would not stay elected for long.

Dangers of Unconstrained Base Money Creation

We should also consider what is likely to happen if helicopter money were implemented. If the policy is perceived to have been successful, then there would be pressure to repeat the operation in order to repeat the “success.” If the policy is perceived to have failed, there would be calls to escalate the program, on the grounds that it only failed because it was not implemented on a grand enough scale.

It gets worse. A new political constituency would have been created—one that realizes that there is “free money” to be had and

all they need to do is lobby for it. From their perspective, the money *would* be free, as they wouldn't bear the costs. The most obvious such constituency would be Congress itself, but one also has to think of the special interest groups that might lobby Congress—that is, everyone who might want “free” money. It is easy to imagine what would happen next. If there were free money to be had, there would be enormous demand for it and there would be corresponding pressure on the Fed to repeat and expand the helicopter money program. The sluice gates would be opened and the remaining constraints against the overissue of base money would be severely, even irremediably, weakened.

Weidmann (2016) also warns of the dangers of uncontrolled monetary activism:

Instead of raising the prospect of ever more daredevil feats, it would actually be wiser to pause for thought. Monetary policy isn't a panacea—it can't replace urgently needed reforms in individual countries, nor can it solve Europe's growth problems. That would simply be too much of a tall order, and it would most certainly end in tears.

Otmar Issing (2016) is even more emphatic. “*Helikopter-Geld ist keine Wunderwaffe,*” he said in a recent interview: “Helicopter money is no wonder weapon.”

I see the entire idea of helicopter money as worrisome, even devastating . . . it is nothing else than a monetary policy declaration of bankruptcy. . . . A central bank that is throwing out money for free, will hardly be able to regain control of the printing-press. . . . I see [this debate] as total intellectual confusion. The economic condition of the world is being meddled into chaos that can't be described [my translation].

Conclusion

Inconvertible fiat currency is unique in that it has a positive exchange value against valuable goods and services, but can be produced at essentially zero cost. The potential temptation for the issuer of fiat currency—the Fed—is to issue unlimited amounts of it, in effect to issue so much of it that its exchange value falls to its

intrinsic commodity value, zero. The fundamental problem in monetary governance is to impose constraints against the overissue of fiat currency in order to preserve its value.

Proposals for helicopter money are the latest in a long line of attempts to kick away constraints against the overissue of fiat currency. They are especially dangerous, however, because they feed the illusion that monetary magic can conjure up real goods and services out of nothing. If this notion takes hold—especially if it takes hold among those who can influence monetary policy, and one thinks especially of Congress—then rational economic policymaking will become impossible. The danger is that the policymaking process will become addicted to helicopter money and that any remaining constraints against the overissue of fiat currency will be thrown out of the window. It is not as if we haven't seen this process play out before. Think of John Law, the Continentals, the Assignats, Weimar, or, more recently, Venezuela or Zimbabwe.

There is a good reason why the issue of paper money has traditionally been subject to tight constraints: it is a dangerous power that once unleashed can easily get out of control. History teaches that we ignore that lesson at our peril, but that from time to time we do so all the same.

Milton Friedman would be horrified at what Keynesians have done to his innocent classroom experiment. John Law, I suspect, would heartily approve.

References

- Bernanke, B. S. (2002) "Deflation: Making Sure 'It' Doesn't Happen Here." Remarks by Governor Ben S. Bernanke before the National Economists Club, Washington (November 21).
- _____. (2015) *The Courage to Act: A Memoir of a Crisis and Its Aftermath*. New York: W. W. Norton.
- _____. (2016) "What Tools Does the Fed Have Left? Part 3: Helicopter Money." *Ben Bernanke's Blog* (April 11). Brookings Institution.
- Bossone, B. (2016) "Why Helicopter Money Is a 'Free Lunch.'" *EconoMonitor* (June 6).
- Brittan, S. (2012) "Come on Bernanke, Fire up the Helicopter Engines." *Financial Times* (August 12).

- Brehon, D., and Winkler, R. (2016) “Helicopters 101: Your Guide to Monetary Financing.” Deutsche Bank Research (April 15).
- Buiter, W. H. (2014) “The Simple Analytics of Helicopter Money: Why It Works—Always.” *Economics* 8 (2014-28): 1–46.
- Buiter, W. H.; Rahbari, E.; and Seydl, J. (2015) “Secular Stagnation: The Time for One-Armed Policy is Over.” Vox, CEPR’s Policy Portal (June 5).
- Caballero, R. (2010) “A Helicopter Drop for the U.S. Treasury.” Vox, CEPR’s Policy Portal (August 30).
- Durden, T. (2013) “Bank of America on the ‘Trillion Dollar Tooth Fairy’ Straight ‘from the Land of Fiscal Make Believe.’” *Zero Hedge* (January 8).
- _____ (2015) “‘The Government Is Literally Paying Itself—Citi Calls for Money Paradrops.’” *Zero Hedge* (September 23).
- Evans-Pritchard, A. (2013) “Helicopter QE Will Never Be Reversed.” *The Daily Telegraph* (April 3).
- Friedman, M. (1968) “The Role of Monetary Policy.” *American Economic Review* 58 (1): 1–17.
- _____ (1969) “The Optimum Quantity of Money.” In M. Friedman, *The Optimum Quantity of Money and Other Essays*, 1–50. Chicago: Aldine.
- Issing, O. (2016) “Ex-EZB-Chefvolkswirt Issing: ‘Helikoptergeld’ Wäre Bankrotterklärung.” *Frankfurter Allgemeine* (March 23).
- Hussman, J. P. (2011) “Charles Plosser and the 50% Contraction in the Fed’s Balance Sheet.” Hussman Funds *Weekly Market Comment* (April 11).
- Hutchinson, M. (2017) “The Bear’s Lair: Mississippi Company Economics.” *True Blue Will Never Stain* (October 17).
- Hutchinson, M., and Dowd K. (2017) “The Apotheosis of the Rentier: How Napoleonic War Finance Kickstarted the Industrial Revolution.” Unpublished manuscript.
- Jordan, J. (2016) “Helicopter Money: Dropping Money from the Sky?” Sound Money Project (April 19).
- Kay, J. (2016) “Essays on Modern Monetary Policy Pt. 1: The Nonsense of Helicopter Money.” JohnKay.com (August 8).
- Keynes, J. M. ([1936] 1973) *The General Theory of Employment, Interest and Money*. Macmillan: London and Basingstoke.
- Krugman, P. (2013) “Be Ready to Mint that Coin.” *New York Times* (January 7).

- Lacalle, D. (2017) “People’s QE: It’s Venezuela with Tea and Cakes.” *Mises Wire* (August 18).
- Muellbauer, J. (2014) “Combatting Eurozone Deflation: QE for the People.” *Vox*, CEPR’s *Policy Portal* (December 23).
- Murphy, R., and Hines, C. (2010) “Green Quantitative Easing: Paying for the Economy We Need.” *Finance for the Future LLP* (December).
- Perrotti, R. (2014) “Eurozone Recovery: There Are No Shortcuts.” *Vox*, CEPR’s *Policy Portal* (September 13).
- Selgin, G. (2016) “Japan: The Way Out.” *Alt-M* (May 3).
- Stockman, D. (2016) “Helicopter Money—the Biggest Fed Power Grab Yet.” *David Stockman’s ContraCorner* (July 14).
- Turner, A. (2015) *Between Debt and the Devil: Money, Credit, and Fixing Global Finance*. Princeton, N.J.: Princeton University Press.
- Weidmann, J. (2016) “Helicopter Money Would Tear Gaping Holes in Central Bank Balance Sheets.” Interview published in the newspapers of the *Funke Mediengruppe* (March 21).
- Wolf, M. (2016) “Helicopter Drops Might Not Be Far Away.” *Financial Times* (February 23).

SORTING OUT MONETARY AND FISCAL POLICIES

Mickey D. Levy

Monetary and fiscal policies have both have gone off track. Excessively easy monetary policy, marked by a massive increase in the Federal Reserve's balance sheet and sustained negative real interest rates, has failed to stimulate faster economic growths but has distorted financial behavior and involves sizable risks. Fiscal policies have resulted in an unhealthy rise in government debt, and projections of dramatic further increases heap burdens on future generations and involve incalculable risks. Monetary and fiscal policies interact in undesirable ways. The Fed's expanded scope of monetary policy has blurred the boundaries with fiscal and credit policies, and the ever-growing government debt may eventually impinge on the Fed and its independence.

A reset of monetary and fiscal policies is required. The Fed has begun to normalize monetary policy so, at this point, a shift in fiscal policy is much more pressing.

The Fed must continue to raise interest rates and unwind its balance sheet but be more aggressive than indicated in its current strategy. The Fed should aim to reduce its balance sheet to the point in which excess reserves are kept relatively low, and it should

Cato Journal, Vol. 38, No. 1 (Winter 2018). Copyright © Cato Institute. All rights reserved.

Mickey D. Levy is Chief Economist for the Americas and Asia at Berenberg Capital Markets, LLC, and a member of the Shadow Open Market Committee. This article is based on testimony presented to the U.S. Congress, House Financial Services Committee, July 20, 2017. The views expressed are the author's and do not reflect those of Berenberg Capital Markets.

fully unwind its holdings of mortgage-backed securities (MBS). A full normalization of monetary policy would benefit economic performance and improve financial health. Equally important, the Fed must acknowledge the limitations of monetary policy and step back from policy overreach, including removing itself from credit allocation policies and toning down its excessive focus on short-term fine-tuning.

The longer-run projections of government debt are alarming and must be taken seriously. General government debt has risen to 100 percent of GDP, up from 61 percent before the 2008–09 financial crisis, while publicly held debt, which excludes debt held for accounting purposes by the Social Security Trust Fund and other trust funds, has risen to 78 percent from 40 percent. The Congressional Budget Office (CBO) estimates that under current law, the publicly held debt-to-GDP ratio is projected to rise to nearly 150 percent by 2047. Congress must develop and implement a strategy that guarantees sound longer-run finances. This requires tough choices, particularly as it addresses the ever-growing entitlement programs, but the costs of inaction are rising. Many acknowledge the risks of rising debt for future economic performance, but in reality the burdens of the government's finances are already affecting current economic performance and the government's allocation of national resources. Witness how the persistent increases in entitlement programs and concerns about high government debt squeeze current spending on infrastructure, research and development, and other activities that would enhance economic performance. Under current laws, these budget constraints—those at the federal level as well as those facing state and municipal governments—will only increase in severity.

Congress's fiscal agenda must be two-pronged. First, Congress must develop and enhance programs and initiatives that directly address the sources of undesired economic and labor market underperformance while restructuring and trimming spending programs that are ineffective and wasteful. This requires transforming the government's annual procedure of budgeting of appropriations for the array of the so-called discretionary programs and dealing with the entitlement programs from a "deficit bean-counting" exercise into a strategic process that carefully assesses the structure of key programs and their objectives—whether they are meeting their policy and social objectives; whether they are doing so effectively; their

unintended side effects; and how they may be enhanced, modified, and cut.

Second, Congress must enact laws that gradually phase in reforms of the entitlement programs that constrain the projected growth of *future* spending in a fair and honest way, improving the benefit structures of the programs with the objectives of protecting lower income retirees and providing sufficient time for older workers to plan for retirement.

The Proper Roles of Monetary and Fiscal Policies

I fully understand the frustrations stemming from the underperformance of the economy in recent years—the sizeable pockets of persistently high unemployment and low wages facing many working-age people, and weak trends in business investment and productivity that underlie disappointingly slow growth. We all want better performance. But the issue is how to achieve it.

Neither the Fed’s sustained monetary ease nor high deficit spending addresses structural challenges facing labor markets, business caution in expansion and investing, weak productivity, and other critical issues. This is particularly apparent with the unemployment rate at 4.3 percent, below standard estimates of its “natural rate” (so-called full employment).

The reality is monetary policy cannot create permanent jobs, improve educational attainment or skills, permanently reduce unemployment of the semi-skilled, or raise productivity and real wages. Rather, monetary policy is an aggregate demand tool. The major sources of underperformance involve structural challenges that are beyond the scope of monetary policy to address. Yet in recent years, there has been excessive reliance on the Fed. All too frequently, analysts and observers opine “fiscal policy is dysfunctional so the Fed has to ease policy.” This assumes that monetary policy and fiscal policy are two interchangeable levers. They are not. Monetary policy is not a substitute for fiscal policy. Monetary policy involves the Fed’s control of interest rates and the amount of money in the economy, which influences aggregate demand and longer-run inflation.

Fiscal policy operates differently. Government spending programs and tax structures allocate national resources—for income support, national defense, health care, public goods like infrastructure, and an array of other activities—and create incentives favoring certain

activities while discouraging others. In a critical sense, the magnitude and mix of spending programs and the structure and details of tax policies—along with the magnitudes of deficit spending—reveal the nation’s priorities set by past and current fiscal policymakers. These allocations of national resources and how specific spending and tax provisions influence households and businesses are key inputs to economic performance, productivity, and potential growth.

In recent decades, the most pronounced change in the federal government’s budget is the rapid expansions of Social Security, Medicare, and Medicaid. According to the CBO (2017), outlays for Social Security, Medicare and Medicaid, and health-care related entitlements have risen from 47 percent of total federal outlays (10.1 percent of GDP) in 1992 to 62.9 percent of federal outlays (13.2 percent of GDP) in 2017. These programs are projected to rise dramatically further to 65.3 percent of federal outlays (15.4 percent of GDP) by 2027.

The objectives of these entitlements are laudable, and they are critical for government and society. However, the growth in these programs has been the primary source of the rising government debt (and projections of further increases), and has significantly increased the share of government spending allocated to income support and health. Consequently, spending on other programs has been squeezed, including those that would enhance longer-run productive capacity. For virtually every state, Medicaid spending is one of the largest and fastest growing spending programs. Faced with rigid balanced-budget constraints on their operating budgets, states have cut back on the provision of some basic government goods and services.

Can these government programs be improved, made more efficient, or modified in ways that maintain their objectives? Yes. Congress must cut through budget categorizations like “mandatory spending” and “discretionary spending programs” and identify ways to improve the efficiency of these programs while maintaining their intent.

Aside from monetary and fiscal policies, labor market performance and business decisions are affected by a growing web of economic and labor regulations imposed by federal, state, and local governments. Private industries add to the list of regulatory requirements, including the expanding imposition of occupational certification requirements and other practices like “noncompete” job contracts. Certainly, while some of these government regulations and

industry rules serve important roles, many constrain the mobility of a sizeable portion of the labor force, limit job opportunities, and are very costly to the economy. Obviously, these are beyond the scope of monetary and fiscal policy.

Regulatory policies deserve attention, in the discussion about the efficacy of monetary and fiscal policies, because they have unique economic effects that may work at cross-purposes to monetary and fiscal policies. In order to establish public policies that improve standards of living, we need to address the sources of economic and labor underperformance with the proper policy tools, rather than rely on standard monetary and fiscal stimulus that are unlikely to have desired outcomes but are costly and generate unintended side effects.

The Fed's Expanded Scope

The Fed deserves credit for its quantitative easing (QE) in 2008–09 that helped to restore financial stability and end the deep recession. The paralysis in the mortgage and short-term funding markets was scary and truly a crisis. The Fed's aggressive interventions and asset purchases, including its large-scale purchases of MBS and its “bailout” of AIG, directly involved the Fed in credit allocation and fiscal policy. At the time, Fed Chairman Ben Bernanke (2008) explicitly identified these Fed interventions as temporary emergency measures, and stated that the Fed would exit them on a timely basis.

But the efficacy of the Fed's unprecedented monetary ease well after the economy had achieved sustainable growth and financial markets had stabilized—the dramatic expansion of its large-scale asset purchase programs (LSAPs) and targeting the Fed funds rate below inflation—is questionable, and the expanded scope of monetary policy involves substantial risks. These policies and the Fed's forward guidance have stimulated financial markets and asset prices, but the economy has been largely unresponsive. Nominal GDP has not accelerated—it has averaged 3.6 percent annualized growth since the Fed implemented QE3 in Fall 2012—and real growth has been subnormal. Business investment has been disappointing despite the Fed's successful efforts to lower the real costs of capital. Productivity gains have been weak, and estimates of potential growth have been reduced significantly. Labor markets have clearly improved, but large pockets of underemployment persist.

Nonmonetary factors including government tax and regulatory policies have hampered credit growth and economic performance (Levy 2017). In banking, the burdensome regulations imposed by Dodd-Frank and the Fed's stress tests have deterred bank lending (Calomiris 2017). The Fed's low rates and forward guidance aimed at keeping bond yields low have dampened expectations. Meanwhile, the Fed's policy of paying interest on excess reserves (IOER), which began in October 2008, at a rate above the effective fed funds rate, has increased the demand to park reserves at the Fed rather than lend them out (Selgin 2016). Despite the dramatic surge in the Fed's monetary base, M2 has grown at a modest 6 percent rate in recent years. The response of aggregate demand has been tepid, and velocity has declined. This reflects several factors. Lower interest rates have increased the demand for money. The Fed's forward guidance has reinforced a sense of caution in the economy. In addition, burdensome regulations have inhibited the supply of bank credit and the monetary policy channels have been clogged. As a result, the high-powered money created by the Fed's LSAPs remain as excess reserves on big bank balance sheets and have not been put to work in the economy (Ireland and Levy 2017).

In the nonfinancial sector, the array of taxes and regulatory burdens and mandated expenses imposed by federal, state, and local governments have constrained business and household spending. Of note, these government-imposed burdens have led businesses to raise their required hurdle rates for investment projects and many job-creating expansion plans have been scuttled. Capital spending has been disappointing in light of the low real costs of capital and strong corporate profits and cash flows.

The Fed takes far too much credit for the sustained economic expansion and labor market improvement of recent years. Without the sustained aggressive monetary ease, the economy would have continued to expand and jobs would have increased. History shows clearly that economic performance has not been harmed when the Fed has normalized interest rates following a period of monetary ease. Not surprisingly, the three Fed rate hikes since December 2015 have had no material impact on economic performance.

Through most of the expansion, the Fed viewed the low wage gains and inflation as a rationale to enhance and maintain its efforts to use monetary policy to stimulate economic growth. Effectively, the Fed's mindset has evolved into the belief that its role is to manage the

real economy. The Fed has rejected the notion that its persistently easy policies have been ineffective—much less the possibility that its policies may have had negative effects. Recently, the Fed has changed its tune. In a speech in March 2017, entitled “From Adding Accommodation to Scaling it Back,” Fed Chair Janet Yellen identified “unwelcome developments” affecting economic performance that “reflect structural challenges that lie substantially beyond the reach of monetary policy.” She continued to state that “Fiscal and regulatory policies—which are of course the responsibility of the Administration and Congress—are best suited to address such adverse structural trends” (Yellen 2017).

Along with weak productivity gains, the failure of nominal GDP to accelerate in response to the Fed’s unprecedented monetary ease has been a key reason why wage increases have remained modest and inflation has remained below the Fed’s 2 percent target. The slow (and nonaccelerating) growth of aggregate product demand has influenced wage and price setting behavior, reducing the flexibility of businesses to raise product prices and reducing their willingness to grant higher wages. Slow growth in nominal GDP—it has averaged 3.4 percent so far this expansion compared to 5.3 percent in the 2001–07 expansion and 5.6 percent during the 1990s—is statistically significant in explaining the slow wage gains despite the low unemployment rate (Levy and Reid 2016). Additionally, inflation has been constrained by declining quality-adjusted prices of select goods and services stemming from technological innovations and product improvements. Most notably, the PCE deflator for durable goods has fallen persistently since the mid-1990s. These innovations have increased consumer purchasing power and benefitted the economy.

The Fed’s historic tendency to fine-tune the economy and financial markets has been accentuated during this expansion. This was apparent when the Fed implemented QE2, Operation Twist, and QE3 in an explicit effort to lower unemployment, and since 2014 when the Fed delayed tapering its asset purchases and then stuck to its reinvestment program to maintain its oversized portfolio. The Fed has been heavily influenced by short-term fluctuations in the economy, and by global and domestic asset markets. It has modified its employment mandate to include focus on the labor force participation rate and wages. These are beyond the Fed’s mandate and well beyond the scope of monetary policy. Such short-term focus and expanded role historically have led to policy mistakes.

The Fed's Balance Sheet

As a result of these short-run concerns, the Fed maintains a balance sheet of \$4.5 trillion, including \$2.5 trillion of U.S. Treasury securities of various maturities and \$1.8 trillion of MBS, primarily with long maturities. The Fed is now the largest holder of each, with 17 percent of outstanding federal publicly held debt and 12 percent of MBS outstanding. (The Fed's holdings of Treasuries are counted as publicly held debt because the Federal Reserve Banks are legally capitalized by the private sector banks in their districts). Prior to the financial crisis, the Fed's balance sheet was roughly \$850 billion, composed nearly entirely of short-term Treasuries and other liquid securities.

The Fed has begun a strategy of gradually and passively unwinding a fairly even portion of its Treasury and MBS holdings by reinvesting all but a small portion of principle of maturing assets. Although the Fed has not been clear about the ultimate size of balance sheet it wishes to maintain, several Fed members have indicated that its ultimate aim is to maintain a large buffer of excess reserves. This strategy should be modified. The Fed's holdings of MBS are inappropriate, directly involving monetary policy in credit allocation, and should be totally unwound. The Fed's MBS holdings effectively favor mortgage credit over other types of credit. While the initial MBS purchases during the height of the financial crisis had a distinct purpose—to stabilize a completely dysfunctional and illiquid market that posed a threat to global markets—continuing to hold MBS makes little sense. Mortgage markets are functioning normally with sufficient liquidity. The Fed's ongoing explicit subsidies of the housing sector are irrational, and the Fed should go back to an all-Treasuries portfolio.

The Fed's intention to maintain a large buffer of excess reserves implies a shift from pre-financial crisis operating procedures, in which the Fed's much smaller asset portfolio resulted in a minimal amount of excess reserves. The Fed has built an argument that maintaining a large amount of excess reserves going forward would benefit the Fed's conduct of monetary policy and enhance its ability to stabilize financial markets. The maintenance of a large buffer of excess reserves would require the Fed to continue to pay IOER and manage the effective Fed funds rate through a "floor system." I prefer a strategy of maintaining a smaller balance sheet that would

involve less excess reserves in the banking system, like the Fed used through most of its history prior to the financial crisis in 2008–09. With minimal excess reserves, the Fed used a market-based “corridor system” in which it managed the effective funds rate close to its target rate through open-market purchases and sales of assets. Going back to its traditional operating procedure would allow the Fed to lessen its exposure in the overnight reverse repo market and in general constrain its overall footprint in financial markets. However, this operational preference is of less importance than the higher priorities of fully winding down the Fed’s MBS holdings and reining in the scope of monetary policy.

Monetary Influences on Fiscal Policy

The Fed’s balance sheet, low policy rate, and forward guidance aimed at keeping bond yields low temporarily have combined to reduce budget deficits and the government’s debt service costs. The Fed effectively is operating a massive positive carry strategy by borrowing short and lending long. This will generate profits and reduce budget deficits as long as interest rates stay low. The Fed’s remittances to the U.S. Treasury reached a peak of \$117 billion in Fiscal Year 2015. They have receded in 2016 and 2017 as the average yield on the Fed’s portfolio has receded and the Fed’s rate hikes have increased the interest it pays to commercial banks under IOER. These large remittances to the Treasury have materially reduced recent budget deficits.

While this deficit reduction may sound good superficially, it involves sizeable risks—to current and future taxpayers—and entangles the Fed’s monetary policy in the government’s budget and fiscal policies in unhealthy ways. Congress seems to perceive that the Fed’s outsized profits remitted to the Treasury are risk-free and permanent, when in fact they involve sizeable interest rate risks. Moreover, the Fed’s balance sheet exposes monetary policy to undesirable budget practices and may undercut the Fed’s independence and credibility.

At a recent congressional hearing held by the House Financial Services Committee on the interaction between monetary and fiscal policies, Congressman Brad Sherman (D–CA) heaped praise on the outsized net profits the Fed remits to the Treasury and asked, “What would the Fed need to do to double (to \$200 billion)

the amount of profits it remits to the Treasury?” This question may seem amusing to monetary economists, but it illustrates a lack of understanding about the proper role of monetary policy and highlights the Fed’s potential vulnerabilities. The Fed should not understate the political-economy risks of maintaining such a large balance sheet. By suppressing deficits and debt service costs, the Fed’s outsized remittances have eased pressure on Congress to address the growing budget imbalance. Also, as illustration of how Congress may misuse the Fed’s large remittances, in December 2015, Congress’s enactment of the FAST Act to provide financing for transportation infrastructure redirected a small portion of the Fed’s assets and some of its net profit into the Highway Trust Fund. The Fed did not protest the way this budgetary “sleight of hand” procedure inappropriately used monetary policy for fiscal purposes.

In light of the magnitude of federal debt outstanding (currently \$15 trillion and estimated by the CBO to rise to \$27 trillion in 2027), budget deficits and debt service costs are very sensitive to interest rates. The CBO (2017) estimates that a 1 percentage point increase in interest rates from its baseline assumptions over the 10-year projection period would add \$1.6 trillion to the budget deficit. Such interest-rate risk must be taken seriously. The Fed’s forecasts of higher policy rates, sustained economic growth, and a rise in inflation to 2 percent point toward higher bond yields. Prior experiences of positive carry strategies often end badly. Witness the failures of many private financial companies, as well as Fannie Mae and Freddie Mac, which required government bailouts. The Fed’s efforts to be more transparent should include a clear and honest assessment of the government’s budgetary risks of its sustained outsized balance sheet.

Fiscal Policy Influences on Monetary Policy

To date the basic thrust of the Fed’s monetary policy has not been materially influenced by budget deficit considerations, although the Fed takes into account fiscal policy deliberations and is sensitive to the impact some of its extraordinary actions have had on the federal budget. But these have been relatively low-level concerns. A much larger concern centers on projections of dramatically rising government debt and the lack of impetus of fiscal policymakers to address

the issue, which raise the prospects that the government's finances may exert burdens on the Fed and impinge on monetary policy.

Sound monetary policy ultimately relies on sound government finances (Leeper 2010). In the extreme, unsustainably high government debt service burdens may dominate monetary policy and require the Fed to accommodate fiscal policy by reducing the real value of the debt or, in an extreme case, ensuring the government's solvency. Such a prospect of fiscal dominance of monetary policy seems remote and far off. However, it may not be so distant, particularly if fiscal policymakers ignore the longer-term budget debt realities. Moreover, nobody really knows when the level of debt becomes "unsustainable" or when or how government finances may unhinge inflationary expectations (Weidmann 2013).

In this context, the current fiscal debate about tax policy should be focusing on reforms that increase productive capacity by reducing inefficiencies and distortions and reducing the disincentives to invest, rather than temporary fiscal stimulus that involves more deficit spending. This is particularly true with the economy in its ninth consecutive year of expansion and clearly displaying signs of self-sustaining growth.

Congress faces several alternative fiscal policy paths. It may continue to avoid reforming current spending programs and the tax structure. This would reinforce disappointing economic performance, and downside risks would rise. Economic growth would remain slow, large pockets of underperformance in labor markets and slow wage growth would persist, reliance on income support would mount and government programs would become increasingly strained, and government debt would continue to rise rapidly. Alternatively, Congress may reform current spending programs, particularly entitlements, by improving their structures while maintaining their intent. Congress also needs to address the sources of the rising government debt and overhaul the tax system. The latter requires reducing marginal tax rates, particularly corporate taxes; broadening the tax base through eliminating the array of deductions, deferrals, exemptions and credits; and simplifying the tax code. Those efforts would lift sustainable economic growth, improve productivity, increase wages and economic well-being of underperformers in labor markets, ease burdens on income support systems, and improve government finances. Future concerns are quickly becoming current realities.

Conclusion

Is it appropriate to be critical of macroeconomic policies amid sustainable economic expansion and low inflation? Yes. Debt projections and the current tax code cry out for reform. The current corporate tax reform initiative is promising. But addressing the government's deficit spending and rising debt is a thorny challenge. How can current fiscal policymakers be expected to make necessary strategic changes to entitlement programs when doing so may involve short-term political fallout? There is no easy answer, but a good starting point would involve members of Congress learning the basic structures and key details of the biggest government spending programs—beneficiary requirements and benefit structures, the magnitude and distribution of benefits, and how they are financed. Such programmatic knowledge could become a healthy basis for a more economically rational policy debate and replace the current tendency of fiscal policymakers to make superficial statements on key programs that only serve to polarize the debate.

Redirecting monetary policy is simpler. The Fed must change its mindset. It should purposely narrow the objective of monetary policy to maintain low inflation and inflationary expectations and be more circumspect about the proper role of monetary policy as an input to sustainable healthy economic performance. This would lead the Fed to redirect itself from its recent thrust of “managing the real economy” and excessive fine-tuning, and acknowledge that some challenges facing the economy are better addressed through other economic, fiscal, and regulatory policy tools. This would steer the Fed away from its harmful roles in fiscal and credit policies.

References

- Bernanke, B. (2008) “Federal Reserve Policies in the Financial Crisis.” Speech at the Greater Austin Chamber of Commerce, Austin, Texas (December 1).
- Calomiris, C. W. (2017) “The Microeconomic Perils of Monetary Policy Experimentation.” *Cato Journal* 37 (1): 1–15.
- CBO (Congressional Budget Office) (2017) *An Update to the Budget and Economic Outlook: 2017 to 2027* (June 29).
- Ireland, P., and Levy, M. (2017) “A Strategy for Normalizing Monetary Policy.” Shadow Open Market Committee (April 21): <http://shadowfed.org/wp-content/uploads/2017.pdf>.

- Leeper, E. (2010) “Monetary Science, Fiscal Alchemy.” Federal Reserve Bank of Kansas City, Jackson Hole Symposium (August 26–28).
- Levy, M. (2017) “Why Have the Fed’s Policies Failed to Stimulate the Economy?” *Cato Journal* 37 (1): 39–45.
- Levy, M., and Reid, R. (2016) “How Much Can Wages Rise?” *Berenberg Capital Markets* (October 26).
- Selgin, G. (2016) “Interest on Reserves and the Fed’s Balance Sheet.” Testimony before the U.S. House of Representatives Committee on Financial Services, Monetary Policy and Trade Subcommittee (May 17). Available at www.cato.org/publications/testimony/interest-reserves-feds-balance-sheet.
- Weidmann, J. (2013) “Who Calls the Shots? The Problem of Fiscal Dominance.” Speech to the BdF-BBk Macroeconomics and Finance Conference, Paris (May 24).
- Yellen, J. (2017) “From Adding Accommodation to Scaling It Back.” Speech to the Executives’ Club of Chicago (March 3).

HEALTH, SUSTAINABILITY, AND THE POLITICAL ECONOMY OF FOOD LABELING

Colin Fristoe

Public choice theory is used to explain regulation of telecommunications, transportation, and other complex production systems (Buchanan 1986). My goal is to show that public choice theory explains a set of regulations that consumers interact with on a daily basis: food labels. I argue that organic labeling serves primarily to protect the organic industry from competition.

“Organic” lacks an unambiguous definition but literally refers to the “farm as organism” (Paull 2006). In the early 20th century, English agriculturalists realized that manufactured nitrogen fertilizer depleted nutrients in soil over time. A movement began to conserve and regenerate soil by encouraging farmers to fertilize with plant residues and animal manure. Today, “organic” has a broader meaning, but the theme is that new technologies undermine the integrity of agriculture. Oregon established the first organic standards program in 1973. Between then and 1990, 21 other states adopted a heterogeneous set of organic standards. The industry believed the multiplicity of standards was causing confusion and undermining the value of the organic label. Through vehicles like the Organic Trade Organization and the American Farm Bureau Federation, the industry petitioned Congress for a set of national organic standards. The Organic Food Production Act of 1990 (OFPA) created the National

Cato Journal, Vol. 38, No. 1 (Winter 2018). Copyright © Cato Institute. All rights reserved.

Colin Fristoe is an undergraduate student at Dartmouth College, studying Economics and Biology. He thanks Professor John Welborn for his guidance.

Organic Standards Board (NOSB) and directed it to propose regulations to the U.S. Department of Agriculture. The USDA finalized its set of rules in 2000, which are periodically amended upon recommendation by the NOSB. (Ellsworth 2001; Pollan 2006; USDA 2000)

Public choice theory is key to explaining the organic standards that emerged in 2000. As Cohrssen and Miller (2016) wrote in *Regulation* magazine, the organic label is “a valuable stamp of approval” from the government, meaning that many people stood to gain from it. No other prospective regulation has received as much input from the public. The USDA was forced to withdraw its first proposal in 1997 after it received an unprecedented 275,603 comments (USDA 2000). The National Organic Program (NOP) is also a striking example of a regulation demanded by its target industry. Consider Secretary of Agriculture Dan Glickman’s comments at the unveiling of the NOP: “The organic label is a marketing tool It is not a statement by the government about food safety. Nor is [it] a value judgment by the government about nutrition or quality” (USDA 2000). Organic labeling is a transparent rent-generating scheme. The question is whether the rents increase social welfare. For this reason, public choice theory is the best way to understand why we have organic standards.

The Organic Standards: Reality and Perception

Like any marketing tool, an organic label is a form of product differentiation. In 2015, organic food sold for an average 47 percent premium (Marks 2015). Clearly, consumers believe it is a better product. But, as Secretary Glickman said, labeling is about managing perceptions, not objective criteria. In this section, I discuss what the organic label means and whether it corrects an externality or information asymmetry problem.

Organic standards are complex but fall into three broad categories: crop standards, livestock standards, and handling standards. Crop production standards prohibit the use of most synthetic pesticides. Synthetic pesticides are organic compounds not produced by a known living organism. Organic pesticides, on the other hand, are typically inorganic metals, like copper and sulfur, or organic plant toxins. Exceptions to the prohibition are granted based on “need” by the NOSB. In addition, organic crops cannot be genetically modified

in a lab, or irradiated, a sterilization process that prolongs shelf life. According to the regulatory text that governs the NOP, published in the Federal Register, organic farms must obtain certification from an accredited private body to confirm that farming practices “maintain or improve the physical, chemical, and biological condition of the soil and minimize soil erosion” (NOP 2011). This determination is made at the discretion of the certifying body. Livestock and poultry producers must meet additional standards. The animals must be fed 100 percent organic feed, and ruminants must receive 30 percent of their dry matter intake from pasture. Animals can’t be given hormones or antibiotics unless sick or malnourished and must have “access to the outdoors.” Finally, handling standards require that *certified organic* products contain 95 percent organic ingredients and prohibit commingling of organic and nonorganic products.

Consumers seem to desire products produced to these standards, yet understand them poorly. The overwhelming majority of consumers purchase organic food out of health concerns, especially perceived deleterious health effects of pesticides, genetically engineered crops, and irradiation (Hughner et al. 2007; Rollin, Kennedy, and Wills 2011). Consumers also believe that organic food tastes better and is more nutritious, although they may simply infer this because it is more expensive. Concerns regarding personal health tend to trump concerns for the environment and animal welfare. However, there is a small subset of consumers that purchases organic food for altruistic reasons as well (Hughner et al. 2007)

The evidence suggests that consumers’ motivations are largely unfounded. Smith-Spangler et al. (2012), in an extensive review in the *Annals of Internal Medicine*, concluded that certified organic products are indistinguishable from conventional ones with respect to safety and nutrition. Consider the public’s primary concern: pesticides. Although there is a perception that organic agriculture is free of pesticides, organic standards allow the use of more than 300 different herbicides and insecticides. Synthetic pesticides, which are not produced by any organism, are generally prohibited, although they may be used when organic pesticides are ineffective. But there is no basis for the distinction between synthetic and natural toxins. In high doses, they exhibit equivalent toxicity and carcinogenicity (Ames, Profet, and Gold 1990). Gold et al. (1992) also found that the potential danger of pesticides is inconsequential amid the background of natural toxins we consume daily. What about the

overapplication of pesticides and supply-chain contamination? Although conventional produce was 30 percent more likely to contain trace pesticides, organic products were just as likely to have pesticide contamination exceeding maximum allowed limits (Smith-Spangler et al. 2012). In the end, both conventional and organic produce are subject to the same consumer safety standards when it comes to chemical residues. Adherence to these standards is the purview of the EPA. Additional regulation might address childhood exposure to areas where pesticides are stored (Muñoz-Quezada et al. 2013), but there is very little evidence that harmful exposure currently occurs through food.

The NOP's absolute prohibition of genetic engineering and irradiation is just as misguided and potentially counterproductive. In fact, when the USDA originally proposed organic standards in 1997, it allowed both methods. The proposal was amended only after a large number of negative public comments (USDA 2000). There is widespread agreement among respected scientific bodies that genetically modified crops are safe for consumption (AMA 2012; AAAS 2012; Key, Ma, and Drake 2008; NAS 2004; WHO 2014). Yet, 39 percent of the public disagrees (Pew Research Center 2016). This poses a problem because GM crop production may have positive externalities. Researchers found, for instance, that on average, GM technology adoption has reduced chemical pesticide use by 37 percent, increased crop yields by 22 percent, and increased farmer profits by 68 percent (Klümper and Qaim 2014). In addition, researchers from the EPA and USDA found that herbicide-resistant GM crops prevent topsoil depletion by eliminating the need for tillage (Fernandez-Cornejo et al. 2012). They also found that GM technology has reduced the amount of pesticides used in farming and the toxicity of these pesticides. Glyphosate, for instance, the herbicide most commonly used with GM crops, is less toxic and not as likely to persist in the environment as the herbicides it replaces (Williams, Kroes, and Munro 2000). Like genetic modification, no available scientific evidence suggests that irradiation is dangerous to human health, and it may actually enhance the safety and quality of food (Bruhn and Wood 1996).

Environmental and animal welfare concerns were the primary impetus for the activists who sought an organic certification program. However, organic farming produces a mixed bag of environmental effects, largely because it shuns promising technologies. I have

already discussed how organic practices may lead to increased pesticide use and topsoil depletion when compared with conventional best practices. Organic yields are also 34 percent lower, meaning that an organic farmer must plant an acre and a half for every acre planted by a conventional farmer to achieve the same yield (Seuvert, Ramankutty, and Foley 2012). One of the ecological arguments for organic farming is that it promotes genetic diversity, but widespread adoption of organic agriculture would devastate hundreds of millions of acres or forestland. New techniques in genetic engineering, such as the CRISPR-Cas9 system, have even greater potential to create designer crops that will yield more food using less land and energy. To be sure, many organic farming practices have benefits. The use of manure, crop rotation, and cover crops leads to higher soil carbon content, which promotes microbial diversity and drought-resistance (Gattinger et al. 2012, Trewavas 2001). Even so, organic production is inefficient compared to an integrated approach that would combine the best practices of both systems.

Similarly, it is not clear that organic livestock standards are effective or objectively improve the lives of animals. Livestock and poultry are required to have *access* to the outdoors. This is the only welfare requirement, and it is up to certifying agencies to interpret. To be fair, the subject is difficult to standardize and involves a number of dilemmas. There is a tradeoff between health, affect, and natural living (Fraser 2008). For instance, organic standards prohibit the use of prophylactic antibiotics, but this leads to higher incidence of parasite-related diseases in organic cattle (Lund and Algers 2003). Some animal welfare advocates approve, wanting to provide as natural a life as possible. Others believe our duty is to spare animals from the agony of disease (Fraser 2008). Intensive systems can provide more shelter and bedding, ration food and water more fairly, and give animals more attention. Not all farmers treat animals well—endemic abuse is well documented. However, organic livestock is treated just as badly as conventional livestock (PETA 2017). Organic food seems to be a luxury good that possesses no advantages over its alternative for humans, animals, or the environment.

By now, I have identified the key organic standards and argued that they cannot be justified by market failure. The organic label does not provide useful information or nudge consumers toward welfare-enhancing products. Because consumers are misled into paying up to 300 percent more, it may even be welfare-reducing. In this case, why

do we have these standards and what would be a better way to address food safety, animal welfare, and the environment? These are the subjects of the next two sections.

Who Benefits?

The USDA has never claimed that organic foods are safer, healthier, or superior to conventional foods in any way. Yet Americans spent \$43 billion on organic products in 2016, and the category has averaged 15 percent growth since 2000 (OTA 2017; FiBL and IFOAM n.d.). How can this be? A representative from the United Kingdom's largest agribusiness consulting firm may have answered that question at the 1999 Organic Food Conference. She said, "If the threats posed by cheaper, conventionally produced products are removed, then the potential to develop organic foods will be limited" (Schroeder et al. 2016). In the years since, the organic industry has erected barriers to entry, stacked its own oversight board, and funded advocacy of dubious science and self-serving legislation. This section examines how and why organic labeling policy is influenced by special interests.

What does the industry have to gain from these activities? The answer should not surprise you: profits. A review of the worldwide organic market in 2014 estimated that organic farmers had 21 percent higher gross returns, 24 percent higher benefit–cost ratios, and 35 percent higher net present values (Crowder and Reganold 2014). Organic farming is also subsidized to the tune of \$160 million, as of the 2014 Farm Bill (USDA 2014). Given the higher net present value, conventional farmers should be rushing into the organic market and driving down the relative price of organic produce. However, significant barriers to entry prevent this from occurring. The most important barriers are the costs of compliance and access to organic markets (Damewood 2015; Strohlic and Sierra 2007). Market access is a barrier because the organic market is only 4 percent of the domestic food market and concentrated. There is a risk that organic producers won't be able to find purchasers (Strohlic and Sierra 2007). Compliance costs include the certification fee, recordkeeping costs, liability, and the three-year transition period during which produce cannot be labeled "organic." Although recordkeeping and inspections costs can amount to thousands of dollars per year, the price premium outweighs these costs. However, violating the regulations may lead to a \$100,000 penalty and imprisonment for up

to five years, consequences that are far more devastating for small, unincorporated producers (Cohrssen and Miller 2016). The transition period imposes a large barrier because farms must absorb lower yields per acre without the offsetting price premium. Like the rest of the law, there is no scientific basis for the three-year transition because common pesticides degrade in weeks or months and need to be sprayed directly onto plants to function. During this period, gross incomes, benefit–cost ratios, and net present values are 10, 7, and 23 percent lower, respectively (Crowder and Reganold 2014). Disease carries an increased risk, because farmers may have to choose between using a pesticide and starting over or losing the entire crop. This requirement is so onerous that some large buyers of organic produce are beginning to offer long-term supply contracts and purchase at a premium during the transition period in order to increase supply (Strom 2015). Whole Foods Inc., the largest buyer of organic produce, tried to create a rating system independent of the organic label that would reduce compliance and liability costs while retaining many of the standards. One supplier complained, “becoming organic is a big investment of time and money and this ratings system kind of devalues all that” (Strom 2015). As George Stigler wrote in 1971, incumbents in regulated industries come to demand regulations. From their point of view, reform would be unfair to those that first ponied up for access to rents.

Organic producers have used regulation to insulate themselves from competition despite controlling only 1 percent of the food market. This is because the organic food market is emblematic of what Mancur Olson (1971) called “concentrated benefits versus diffuse costs.” I have discussed how organic producers benefit, but there is also a subset of organic consumers who strongly *believe* that they benefit from organic certification. Advertisers call these people the “LOHAS” and estimate that they make up 18 percent of their market (Schroeder et al. 2016). LOHAS are “early adopters and influencers” of policy. At the other end of the spectrum, the group least likely to support organic labeling is the “unconcerneds,” who by their very nature do not exhibit strong preferences one way or the other. Conventional producers may also have weak preferences regarding labeling. So far, there is no literature on the effect of OFPA on conventional food prices, but the organic market may still be too niche to affect the world market for produce. The American Farm Bureau Federation supports or tolerates organic labeling, even as it strongly

opposes GMO labeling (Duvall 2017). From its point of view, OFPA benefits 1 percent of farmers without hurting the other 99 percent. The costs instead fall on consumers.

Players in the organic market influence policymaking in several ways. One common mechanism is to fund advocacy groups that lobby for legislation that would help the industry. Table 1 identifies prominent donors from the organic industry that support such groups. Advocacy groups typically seek to influence public opinion by funding and marketing research that portrays conventional agriculture as dangerous. For example, one researcher published a study shortly after Stanford researchers Smith-Spangler et al. (2012), claiming to have found “a 94 percent reduction in health risk from the selection of organic brands.” As phrased, the statement is almost nonsensical and cannot possibly be accurate given the complexity of estimating environmental risk over a lifetime. It didn’t matter that the researcher was a lead scientist for the Organic Trade Organization and the publication was only meant to obscure the issue and undermine the Stanford study (Benbrook 2012; Schroeder et al. 2016). Companies in the industry also join together to file lawsuits and organize lobbying campaigns for legislation against conventional or biotech agriculture. Anti-GMO labeling campaigns are particularly common (Schroeder et al. 2016). Studies have shown that food labels easily influence consumers due to what is known as “framing” (Levin and Gaeth 1988). Organic labels trick consumers into believing that organic food is healthier, safer, and tastier; and mandatory GMO labeling would enhance this perception. The industry is also engaged in direct lobbying of policymakers, as noted by Henry Miller and Julie Kelly (2015). They found that the Environmental Working Group, which advocates against biotechnology and receives donations from 20 organic producers, spent \$1.4 million to lobby U.S. House and Senate members in 2013 and 2014. During the same time period, the Center for Food Safety spent \$1.1 million lobbying, and since 2008 the Organic Trade Association has also spent more than \$1 million on lobbyists. The Organic Consumers Association and Organic Trade Organization, both industry groups, have their own political action committees (PACs) for campaigning on behalf of politicians. Individuals in the industry also contribute money to the political process. One example is the CEO of Stonyfield Farms, who contributed \$419,000 to political campaigns between 2008 and 2014 and was a “bundler” for President Barack Obama. Whole Foods

TABLE 1
PLAYERS IN THE ORGANIC MARKET

Anti-GMO/Pesticide Advocacy Organization	Funders and Campaign Supporters
Beyond Pesticides (2012 budget \$1.1 million)	Amy's Kitchen, Horizon Organic, Earthbound Farms, Stonyfield Organic, Choice Organic Teas, Whole Foods Market, United Natural Foods, Organic Valley, Vitamin Cottage Natural Foods, Annie's Homegrown, Rudi's Organic, Nature's Path, Kamut International, Oregon Tilth, National Cooperative Grocers Association, Dr. Bronner's, Frey Organic Vineyards, Good Earth Natural Foods, Organic Foods International, Demeter Association, Lundberg Family Farms, Organic Seed Growers and Trade Association
Center for Food Safety (2011 budget \$2.9 million)	Horizon Organic, Earthbound Farms, Stonyfield Organic, Whole Foods Markets, United Natural Foods, Organic Valley, Amy's Kitchen, New Chapter Organics, Annie's Homegrown, Big Carrot Natural Foods, Lundberg Family Farms, Eatmore Sprouts and Greens, Organic Seed Growers and Trade Association
Cornucopia Institute (2012 budget \$669,000)	Organic Valley, Frey Organic Vineyards, Demeter Association, Organic Seed Growers and Trade Association
Environmental Working Group (2011 budget \$5.9 million)	Earthbound Farm, Stonyfield Organic, Whole Foods Market, Organic Valley, Applegate Organic, Lundberg Family Farms
Food and Water Watch (2011 budget \$11.1 million)	Stonyfield Organic, Organic Valley, National Cooperative Grocers Association

(Continued)

TABLE 1 (*Continued*)
PLAYERS IN THE ORGANIC MARKET

Anti-GMO/Pesticide Advocacy Organization	Funders and Campaign Supporters
Institute for Agriculture Trade Policy (2012 budget \$3.9 million)	United Natural Foods, Organic Valley, Nature's Path Food, Dr. Bronner's, National Cooperative Growers Association, Good Earth Natural Foods, Organic Seed Growers and Trade Association, Nature's Path Food, Lundberg Family Farms, Organic Valley, Natural Grocer Company, New Hope Natural Media
Non-GMO Project (2011 budget \$455,000)	Stonyfield Organic, Whole Foods Market, United Natural Foods, White Wave Foods, Organic Valley, Nature's Path Food, Big Carrot Natural Foods, Dr. Bronner's, National Cooperative Grocers Association, Lundberg Family Farms, Natural Grocery Company, New Hope Natural Media
The Organic Center (2012 budget \$728,000)	Organic Trade Association, Horizon Organic, Stonyfield Organic, Whole Foods Market, United Natural Foods, White Wave Foods, Organic Valley, Rudi's Organic, Annie's Homegrown, Nature's Path Food, Kamut International, Lundberg Family Farms
Organic Consumers Association (2011 budget \$2 million)	Attune Foods, United Natural Foods, Organic Valley, Amy's Kitchen, Nature's Path Foods, Traditional Medicinals, Dr. Bronner's, Demeter Association, Lundberg Family Farms, Stonyfield Organic
Pesticide Action Network (2011 budget \$2.3 million)	Stonyfield Organic, Organic Valley

TABLE 1 (*Continued*)
PLAYERS IN THE ORGANIC MARKET

Anti-GMO/Pesticide Advocacy Organization	Funders and Campaign Supporters
Sierra Club (2011 budget \$97.8 million)	Stonyfield Organic, Whole Foods Market, Organic Valley, Frey Organic Vineyards
Just Label It (Organic Industry Coalition formed in 2011 that has not yet reported expenditures)	Stonyfield Organic (founder), Horizon Organic, Attune Foods, Earthbound Farm, Silk Soymilk, United Natural Foods, Organic Valley, KeHe, Amy's Kitchen, New Chapter Organic, Frontier National Products, Rudi's Organic, Nature's Path, Annie's Homegrown, Wholesome Sweeteners, Pure Brands, Kamut International, Mom's Organic, Dr. Bronner's, National Cooperative Grocers Association, Earth Source, Good Earth Natural Foods, Organic Foods International, Demeter Association, Lundberg Family Farms, Country Choice Organic, New Hope Natural Media, SunFood, Country Choice
Rodale Institute (2011 budget \$3.7 million)	Earthbound Farms, Attune Foods, Stonyfield Organic, Whole Foods Market, United Natural Foods, Organic Valley, Amy's Kitchen, Frontier National Products, Rudi's Organic, Nature's Path, Traditional Medicinals, Wholesome Sweeteners, Lundberg Family Farms

SOURCE: Adapted from Schroeder et al. (2016: 20).

co-CEO Walter Robb also contributed \$164,000 during the same period. This is only a sample of what Miller and Kelly (2015) call the “money trail,” but it establishes that lobbying and influence peddling are common in the organic food industry.

The industry also exerts influence through its oversight board, the National Organic Standards Board. The NOSB is responsible for compiling the list of allowed and prohibited substances, issuing exemptions, and recommending new rules and amendments. The 15-member board ostensibly reserves seats for organic producers, processors, and retailers, as well as environmentalists, consumer advocates, a certifying agent, and a scientist. In reality, I found that 13 of 15 current members previously worked for, owned, or served on the board of directors of corporations in the industry, trade groups, or advocacy organizations funded by the industry. In the past, representatives for Whole Foods (twice), Driscoll’s, Campbell’s Soup, Earthbound Farm, Horizon Organic, and Smucker’s held seats on the NOSB while concurrently working as organic compliance officers at their companies. The “scientists” on the board worked for General Mills, Earthbound Farms, or California Certified Organic Farmers, a trade group. “Consumer/Public Interest” advocates and “environmentalists” have frequently included marketing consultants for organic producers, representatives of organic advocacy organizations, and owners of companies that qualify as organic producers, handlers, or retailers. Realistically, the only people with an incentive to serve on the NOSB work in the organic industry and benefit from organic standards. That is why even “consumer advocates” tend to have vested interests in maintaining food-labeling regulation. The NOSB is no exception to the revolving-door phenomenon.

Through the NOSB, direct lobbying, and public opinion campaigns, the organic industry maintains a profit-generating regulatory fiefdom. Using public choice theory, this explains why the National Organic Program exists despite the lack of compelling evidence that it corrects a market failure. The next section will focus on how a better system might work, and the political hurdles of implementing one.

Conclusion

I take the view that the National Organics Program should be repealed. That does not mean that information asymmetries do not

exist in the food market, but that the NOP does not fix them. The first-best option is a system that would convey perfect information to the consumer at the point of purchase. Since that is not available, we have to choose between a government standard and a market-based labeling scheme. A government standard provides an opportunity for rent-seeking and regulatory capture. The disadvantage of a market-based scheme is that quality is difficult to evaluate. A multiplicity of different criteria will likely confuse consumers and undermine the value of the label to producers. However, concerned consumers would rely on discovery, and private firms would have an incentive to compete for the best criteria. Examples of private food labels already exist and include Oregon Tilth, Free-Range, Grass Fed, and Animal Welfare Approved.

The argument for keeping the status quo notes that the NOP is completely voluntary. No one forces organic standards on farmers or organic products on consumers. Those customers seem to gain some utility from the USDA label. Who are we to evaluate their purchases for them? This argument assumes rationality, but we know that consumers can be “nudged” in subconscious ways. If they knew that the standards did not address safety, they might resent this kind of manipulation. And for those who completely disagree with my assessment of the standards, private labels like Oregon Tilth could step in.

The USDA should never have engaged in marketing on behalf of a small market segment. It may have been a rational response to the political landscape, but it is cynically redistributive and distorts market incentives. The NOP should be repealed but likely will not be because the benefits of doing so would be more diffuse than the costs.

References

- American Association for the Advancement of Science (AAAS) (2012) “Statement by the AAAS Board of Directors on Labeling of Genetically Modified Foods” (October 20). Available at www.aaas.org/sites/default/files/AAAS_GM_statement.pdf.
- American Medical Association (AMA) (2012) “Report of the Council on Science and Public Health: Labeling of Bioengineered Foods.” Available at www.ama-assn.org/sites/default/files/media-browser/public/about-ama/councils/Council%20Reports/council-on-science-public-health/a12-csaph2-bioengineeredfoods.pdf.

- Ames, B. N.; Profet, M.; and Gold, L. S. (1990) "Nature's Chemical and Synthetic Chemicals: A Comparative Toxicology." *Proceedings of the National Academy of Sciences* 87 (19): 7782–86.
- Benbrook, C. M. (2012) "Impacts of Genetically Engineered Crops on Pesticide Use in the U.S.: The First Sixteen Years." *Environmental Sciences Europe* 24: 24–36.
- Bruhn, C. M., and Wood, O. B. (1996) "Position of the American Dietetic Association: Food Irradiation." *Journal of the American Dietetic Association* 96 (1): 246–53.
- Buchanan, J. (1986) "The Constitution of Economic Policy, Nobel Prize Lecture." *American Economic Review* 77 (3): 243–50.
- Cohrssen, J. J., and Miller, H. I. (2016) "The USDA's Meaningless Organic Label." *Regulation* 39 (1): 24–27.
- Crowder, D. W., and Reganold, J. P. (2014) "Financial Competitiveness of Organic Agriculture on a Global Scale." *Proceedings of the National Academy of Sciences* 112 (24): 7611–16.
- Damewood, K. (2015) *Report on Economic Barriers to Organic Transition*. Santa Cruz: California Certified Organic Farmers (CCOF).
- Duvall, Z. (2017) "Annual Address to Members of the American Farm Bureau Federation." Available at www.fb.org/newsroom/farm-bureau-president-to-congress-slash-regulations-reform-immigration-ag-1.
- Ellsworth, J. (2001) "The History of Organic Food Regulation." *HLS Student Papers*. Cambridge, Mass.: Harvard University Press.
- Fernandez-Cornejo, J.; Hallahan, C.; Nehring, R.; Wechsler, S.; and Grube, A. (2012) "Conservation Tillage, Herbicide Use, and Genetically Engineered Crops in the United States: The Case of Soybeans." *Journal of Agrobiotechnology Management & Economics* 15 (3): 231–41.
- Forschungsinstitut für biologischen Landbau (FiBL, Research Institute of Organic Agriculture) and International Federation of Organic Agriculture Movements (IFOAM) (n.d.) "Organic Food Sales Growth in the United States from 2000 to 2015." *In Statista: The Statistics Portal*. Available at www.statista.com/statistics/196962/organic-food-sales-growth-in-the-us-since-2000.
- Fraser, D. (2008) "Understanding Animal Welfare." *Acta Veterinaria Scandinavica* 50 (1): S1.

- Gattinger, A.; Muller, A.; Haeni, M.; Skinner, C.; Fliessbach, A.; Buchmann, N.; Mäder, P.; Stolze, M.; Smith, P.; El-Hage Scialabba, N.; and Niggli, U. (2012) “Enhanced Top Soil Carbon Stocks under Organic Farming.” *Proceedings of the National Academy of Sciences* 109 (44): 18226–231.
- Gold, L. S.; Sloan, T. H.; Stern, B. R.; Manley, N. B.; and Ames, B. N. (1992) “Rodent Carcinogens: Setting Priorities.” *Science* 258 (5080): 261–65.
- Hughner, R. S.; McDonagh, P.; Prothero, A.; Shultz, C. J.; and Stanton, J. (2007) “Who Are Organic Food Consumers? A Compilation and Review of Why People Purchase Organic Food.” *Journal of Consumer Behaviour* 6 (2–3): 94–110.
- Key, S.; Ma, J. K. C.; and Drake, P. M. W. (2008) “Genetically Modified Plants and Human Health.” *Journal of the Royal Society of Medicine* 101 (6): 290–98.
- Klümper, W., and Qaim, M. (2014) “A Meta-Analysis of the Impacts of Genetically Modified Crops.” *PLoS ONE* 9 (11): e111629.
- Levin, I. P., and Gaeth, G. J., (1988) “How Consumers Are Affected by the Framing of Attribute Information Before and After Consuming the Product.” *Journal of Consumer Research* 15 (3): 374–78.
- Lund, V., and Algers, B. (2003) “Research on Animal Health and Welfare in Organic Farming: A Literature Review.” *Livestock Production Science* 80 (1–2): 55–68.
- Marks, T. (2015) “The Cost of Organic Food.” *Consumer Reports* (March 19).
- Miller, H. I., and Kelly, J. (2015) “Government Favors and Subsidies to Organic Agriculture: Follow the Money.” *Forbes* (September 23).
- Muñoz-Quezada, M. T.; Lucero, B.A.; Barr, D. B.; Steenland, K.; Levy, K.; Ryan, P. B.; Iglesias, V.; Alvarado, S.; Concha, C.; Rojas, E.; and Vega, C. (2013) “Neurodevelopmental Effects in Children Associated with Exposure to Organophosphate Pesticides: A Systematic Review.” *NeuroToxicology* 39: 158–68.
- National Academy of Sciences (NAS) (2004) *Safety of Genetically Engineered Foods: Approaches to Assessing Unintended Health Effects*. Washington: National Academies Press.
- National Organic Program (NOP) (2011) 7 C.F.R. § 205. Washington: Government Publishing Office.

- Olson, M. (1971) *The Logic of Collective Action: Public Goods and the Theory of Groups*. Cambridge, Mass.: Harvard University Press.
- Organic Trade Association (OTA) (2017) “2016 Organic Industry Survey” (conducted January 7–March 25). Available at www.ota.com/news/press-releases/19031.
- Paull, J. (2006) “The Farm as Organism: The Foundational Idea of Organic Agriculture.” *Elementals: Journal of Bio-Dynamics Tasmania* 80: 14–18.
- People for the Ethical Treatment of Animals (PETA) (2017) “Free-Range and Organic Meat, Eggs, and Dairy Products: Conning Consumers?” Available at PETA.org/issues/animals-used-for-food/free-range-organic-meat-myths.
- Pew Research Center (2016) “The New Food Fights: U.S. Public Divides over Food Science.” Survey (conducted May 10–June 6). Available at www.pewinternet.org/2016/12/01/the-new-food-fights.
- Pollan, M. (2006) *The Omnivore’s Dilemma: A Natural History of Four Meals*. London: Penguin Press.
- Rollin, F.; Kennedy, J.; and Wills, J. (2011) “Consumers and New Food Technologies.” *Trends in Food Science and Technology* 22 (2): 99–111.
- Schroeder, J.; Chassy, B.; Tribe, D.; Brookes, G.; and Kershen, D. (2016) “Organic Marketing Report.” *Academics Review*. Available at http://academicsreview.org/wp-content/uploads/2014/04/Academics-Review_Organic-Marketing-Report1.pdf.
- Seuvert, V.; Ramankutty N.; and Foley, J. A. (2012) “Comparing the Yields of Organic and Conventional Agriculture.” *Nature* 485: 229–32.
- Smith-Spangler, C.; Brandeau, M. L.; Hunter, G. E.; Bavinger, J. C.; Pearson, M.; and Eschbach, P. J. (2012) “Are Organic Foods Safer or Healthier than Conventional Alternatives?: A Systematic Review.” *Annals of Internal Medicine* 157: 348–66.
- Stigler, G. (1971) “The Theory of Economic Regulation.” *Bell Journal of Economics and Management Science* 2 (1): 3–21.
- Strom, S. (2015) “Recalls of Organic Foods on the Rise, Report Says.” *New York Times* (August 20).
- Strochlic, R., and Sierra, L. (2007) “Conventional, Mixed and ‘Deregistered’ Farmers: Entry Barriers and Reasons for Exiting Organic Production in California.” *California Institute for Rural Studies*. Available at www.cirsinc.org/publications.

- Trewavas, A. (2001) "Urban Myths of Organic Farming." *Nature* 410: 409–10.
- U.S. Department of Agriculture (USDA) (2000) "National Organic Standards: Remarks Prepared for Delivery by Secretary of Agriculture Dan Glickman." Press Release. Available at www.iafp.org/files/National_Organic_Standards_Remarks.htm.
- _____ (2014) *The Agricultural Act of 2014: Highlights and Implications*. Washington: Economic Research Service. Available at www.ers.usda.gov/agricultural-act-of-014-highlights-and-implications.
- Williams, G. M.; Kroes, R.; and Munro, I. C. (2000) "Safety Evaluation and Risk Assessment of the Herbicide Roundup and Its Active Ingredient, Glyphosate, for Humans." *Regulatory Toxicology and Pharmacology* 31 (2): 117–65.
- World Health Organization (WHO) (2014) "Frequently Asked Questions on Genetically Modified Foods." Available at www.who.int/foodsafety/areas_work/food-technology/faq-genetically-modified-food/en.

A REVIEW OF THE REGIONAL GREENHOUSE GAS INITIATIVE

David T. Stevenson

The nearly decade-old Regional Greenhouse Gas Initiative (RGGI) was always meant to be a model for a national program to reduce power plant carbon dioxide (CO₂) emissions. The Environmental Protection Agency (EPA) explicitly cited it in this fashion in its now-stayed Clean Power Plan. Although the RGGI is often called a “cap and trade” program, its effect is the same as a direct tax or fee on emissions because RGGI allowance costs are passed on from electric generators to distribution companies to consumers. More recently, an influential group of former cabinet officials, known as the “Climate Leadership Council,” has recommended a direct tax on CO₂ emissions (Shultz and Summers 2017).

Positive RGGI program reviews have been from RGGI, Inc. (the program administrator) and the Acadia Center, which advocates for reduced emissions (see Stutt, Shattuck, and Kumar 2015). In this article, I investigate whether reported reductions in CO₂ emissions from electric power plants, along with associated gains in health benefits and other claims, were actually achieved by the RGGI program. Based on my findings, any form

Cato Journal, Vol. 38, No. 1 (Winter 2018). Copyright © Cato Institute. All rights reserved.

David Stevenson is Director of the Center for Energy Competitiveness at the Caesar Rodney Institute. He prepared this working paper for Cato’s Center for the Study of Science. He thanks Pat Michaels and Jim Dorn for helpful comments on earlier drafts.

of carbon tax is not the policy to accomplish emission reductions. The key results are:

- There were no added emissions reductions or associated health benefits from the RGGI program.
- Spending of RGGI revenue on energy efficiency, wind, solar power, and low-income fuel assistance had minimal impact.
- RGGI allowance costs added to already high regional electric bills. The combined pricing impact resulted in a 12 percent drop in goods production and a 34 percent drop in the production of energy-intensive goods. Comparison states increased goods production by 20 percent and lost only 5 percent of energy-intensive manufacturing. Power imports from other states increased from 8 percent to 17 percent.

The regional program shifted jobs to other states. A national carbon tax would shift jobs to other countries. A better policy to reduce CO₂ emissions is to encourage innovation rather than rely on taxes and regulation. The United States has already reduced emissions 12 percent from 2005 to 2015, more than any other developed country with a large economy, mainly through innovations in natural gas drilling techniques. There are many other opportunities to invest in innovation, for example, improved solar photovoltaic cells, more efficient batteries, small modular nuclear reactors, and nascent technologies that use fossil fuels without emitting CO₂.

Background

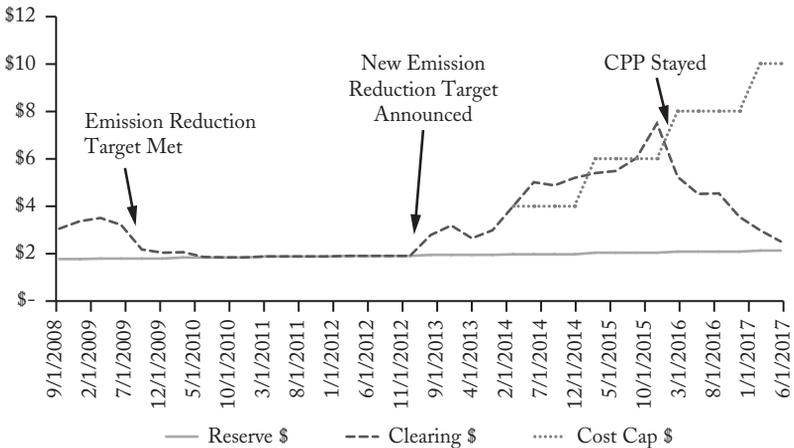
Ten northeast states joined together to form the RGGI to require power plants with a capacity of more than 25 megawatts to buy emission allowances for each ton of CO₂ emissions. The states included Connecticut, Delaware, Maine, Maryland, Massachusetts, New Hampshire, New Jersey, New York, Rhode Island, and Vermont. The allowances were sold in quarterly auctions beginning in 2008. The initial plan was to gradually reduce the number of allowances available to achieve a 10 percent emission reduction by 2018. New Jersey dropped out of the plan in 2011. In 2013, RGGI, Inc. announced plans for a 45 percent reduction in the number of allowances available in auctions beginning in 2014, with an additional 2.5 percent reduction each year until 2020 (Brown 2013: 1).

REGIONAL GREENHOUSE GAS INITIATIVE

Consequently, allowance prices began to rise, and RGGI states are now negotiating an extension to 2030, with an additional 30 percent reduction in allowable emissions.

The program is touted by RGGI, Inc. as a market-based system. However, the program applies a minimum reserve price and a Cost Cap Reserve that kicks in additional allowances if an annual price cap is exceeded (Figure 1). The proposed agreement for 2030 also includes an Emissions Containment Reserve whereby states can withhold allowances if auction prices fall below a set target price. A true market-based cap and trade program would allow the market to set the price. Allowance prices averaged about \$3/ton from 2008 to 2013 ranging from about \$2 to \$4. In 2014, there was a dramatic cut in the number of available allowances that forced prices to a high of \$7.50/ton in 2015, tracking the Cost Cap Reserve target. Prices began to fall after the Clean Power Plan implementation was stayed by the Supreme Court, and hit \$2.53/ton in June, 2017, compared to a reserve price of \$2.15. The extension targets a \$13/ton price in 2021, and \$24/ton in 2030. Speculators have made up roughly one-quarter of allowance purchases, trading with compliance entities in a secondary market.

FIGURE 1
RGGI ALLOWANCE PRICE AUCTION HISTORY



SOURCE: RGGI, Inc., Auction Results.

According to Hibbard et al. (2011:15), in a report for the Analysis Group, “Within the electric system, the impacts of these initial (RGGI) auctions show up during the 2009–11 period, as power plant owners priced the value of CO₂ allowances into prices they bid in regional wholesale prices.” A flow diagram in that report (p. 22) shows how the auction costs flow from the electric generators to the electric distributors, and on to consumers, the same as a direct tax or fee would do.

In order to claim success for RGGI, the first cap and trade program in the United States, we need to consider some related issues:

1. Can the measured emission reductions be accounted for by non-RGGI causes?
2. Can the impacts on the economy be clearly broken down into statistically confirmable independent (RGGI inputs) and dependent variables (real GDP, or electric price changes)?
3. Can the RGGI revenue expenditures be shown to have been necessary and to have had significant impacts?
4. Were energy efficiency project claimed savings rigorously tested by weather-adjusted “before and after” meter readings?

RGGI fails to answer these questions. Unfortunately, the data needed for a robust statistical analysis (question 2) are not readily available and obtaining them is beyond the scope of this article. The other three are noted in the text that follows.

Electricity Demand

The change in electricity demand, by necessity, must consider the interplay of real economic growth, the details of that growth, changes in population, the impact of pricing, and changes in energy efficiency. The RGGI program has an impact on these parameters.

It is difficult to compare electric prices from state to state because of significant regional differences in power cost. Also, at roughly the same time RGGI started, many states began requiring increased use of energy sources like wind and solar in their Renewable Portfolio Standard (RPS) laws and set energy efficiency requirements.

A further complication is that a number of states deregulated the supply portion of electric bills allowing market competition just prior

to the start of the RGGI program. All the RGGI states deregulated. Fortunately, there is a comparison sample of five non-RGGI states (Illinois, Ohio, Oregon, Pennsylvania, and Texas) that deregulated electric supply in a manner similar to the RGGI states, and also had significant RPS requirements. Both RGGI and non-RGGI states have wide variation in their RPS programs, which adds uncertainty. Increasing wind and solar power raises electric rates because they are premium-priced power sources. For example, the increase in Delaware's electric prices by 9 percent is directly related to the RPS, which shows up on consumers' Delmarva Power electric bills. Likewise, Maryland electric bills have increased 14 percent for the same reason, according to a report from the Maryland Energy Administration (Tung 2017: 17).¹

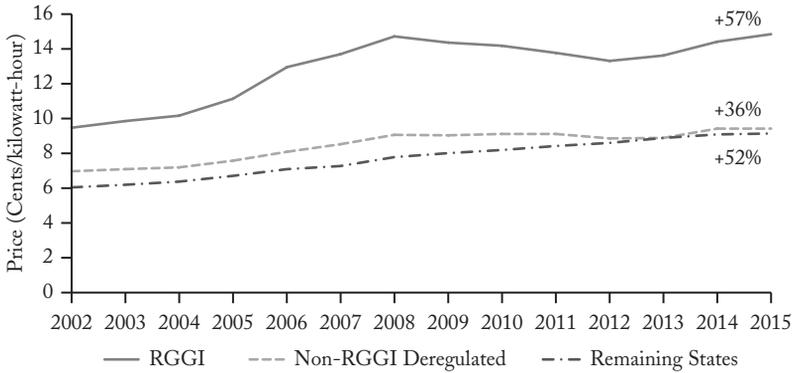
Non-RGGI comparison states actually added more wind and solar generation than RGGI states: adding 5.5 percentage points to generation compared to 2.3 percentage points in the RGGI states. Even removing the large wind farm construction effort in Texas from the calculation, the non-RGGI comparison states still outperformed the RGGI states: adding 3.4 percentage points compared to 2.3. The cost of wind and solar power has averaged two to three times the megawatt-hour rate compared to existing conventional fuel sources. The price impact should be greater in the non-RGGI states. Despite this disadvantage, the non-RGGI states still had lower overall price increases.

Several states that offered limited deregulation were not included in the comparison, and New Jersey is not included as an RGGI state because it dropped out of RGGI in 2011, and California is not included because it began a carbon tax just a few years ago. The results shown in Figure 2 cover the period from 2002 to 2015 to capture the impact of the four policies taken together (deregulation, RPS, energy efficiency, and RGGI).

To more accurately isolate the impact of RGGI between 2007 and 2015, the weighted average of total electric revenue for the multistate groups is used in Table 1, and shows RGGI prices rose

¹I use 2007 as the base year through 2015, unless otherwise noted. The reason for using 2007 is that RGGI auctions began in 2008, which was also the first year of the Great Recession.

FIGURE 2
ELECTRICITY PRICE TRENDS, 2002–15



SOURCE: U.S. EIA (2016a) Detailed State Electricity Data, 1990 to 2015.

64 percent more than comparison states. The increase was split between direct RGGI cost pass-through and indirect cost. Direct emission allowance cost was \$436 million in 2015, about half the price differential between RGGI and comparison states. The rest of the difference may be due to indirect RGGI costs. For example, when power is imported to Delaware and Maryland from the PJM Regional Transmission Organization, there are premium charges for transmission distances, transmission congestion, and

TABLE 1
WEIGHTED AVERAGE NOMINAL PRICE CHANGE,
2007 vs. 2015

	Electric Revenue \$Billion 2007	Electric Revenue \$Billion 2015	Demand Million MWh 2007	Demand Million MWh 2015	\$/MWh 2007	\$/MWh 2015	% Change
RGGI	\$50.8	\$51.4	353.7	342.4	\$143.51	\$150.14	+4.6%
Non-RGGI	\$77.1	\$81.3	851.9	873.8	\$90.51	\$93.01	+2.8%
U.S.	\$343.7	\$391.3	3,482	3,759	\$98.71	\$104.11	+5.5%

SOURCE: U.S. EIA (2016a) Detailed State Electricity Data, 1990 to 2015.

capacity. An earlier study, “Cost Impacts of 2013 RGGI Rule Changes in Delaware” (Stevenson and Stapleford 2016: 2), demonstrates that RGGI allowances directly added \$11 million a year to Delaware electric bills, while the indirect costs added another \$28.5 million.

Prices in RGGI states rose concurrent with more energy-intense manufacturing segments of the economy leaving the RGGI states with slower overall real economic growth based on Regional Real Chained GDP (Table 2). Linking real economic growth to RGGI alone is fraught with problems: real economic growth rates in RGGI states between 2007 and 2015 varied widely from a negative 7.1 percent for Connecticut to a plus 11.9 percent for Massachusetts. Can we realistically claim RGGI helped Massachusetts but hurt Connecticut at the same time?

The comparison states’ economies grew 2.4 times faster than the RGGI states. Data from the U.S. Bureau of Economic Analysis show that the RGGI states lost 34 percent of energy-intensive businesses (primary metals, food processing, paper products, petroleum refining, and chemicals); the comparison states lost only 5 percent. The RGGI states lost 12 percent of overall goods production, while the comparison states grew by over 20 percent. We see this impact show up in industrial electric demand with the RGGI states falling 18 percent, while non-RGGI comparison states fell only 4 percent (Table 3).

Consideration also needs to be given to energy efficiency improvements as shown by the improvement in energy intensity (Table 4). RGGI states improved by 9.6 percent, while non-RGGI comparison states improved 11.5 percent. (Energy intensity improves when it goes down.)

According to RGGI, Inc. (2016), RGGI states are investing the RGGI revenue in energy efficiency projects, suggesting RGGI states should be improving energy efficiency faster than other states. Based on gains in overall energy intensity, this claim appears to be false. An explanation for this disparity may be that the funds are not going to energy efficiency, or that the energy efficiency projects may not be working well. Both effects are seen in Delaware where 35 percent of allowance revenue is assigned to the Department of Natural Resources & Environmental Control (DNREC), and the rest flows through a private, nonprofit organization known as the Sustainable Energy Utility (SEU). Delaware has received \$100 million in RGGI

TABLE 2
REAL GDP GROWTH, 2007–15 (Billions of 2009 Dollars)

State	Energy-Intensive Goods			Total Goods			Total GDP			Change
	2007	2015	2007	2015	2007	2015	2007	2015		
Connecticut	22.2	6.1	50.9	32.6	246.1	228.5			-7.1%	
Delaware	4.1	2.2	8.3	6.5	57.6	60.5			5.0%	
Maine	2.7	1.7	9.1	7.3	51.5	51.1			-0.8%	
Maryland	9.0	7.9	35.8	33.3	301.3	322.0			6.9%	
Massachusetts	11.8	11.8	61.3	59.9	391.0	437.6			11.9%	
New Hampshire	1.2	1.1	10.6	9.9	62.4	64.9			4.2%	
New York	36.1	26.5	117.1	107.1	1,153.5	1,265.6			9.7%	
Rhode Island	1.1	1.3	7.1	6.2	49.7	49.7			0.0%	
Vermont	<u>0.8</u>	<u>0.6</u>	<u>4.7</u>	<u>4.2</u>	<u>25.8</u>	<u>27.2</u>			<u>5.7%</u>	
RGGI Total	89.1	59.2	304.7	267.0	2,338.8	2,507.3			7.2%	
Illinois	32.4	36.7	125.9	120.5	671.1	686.0			2.2%	
Ohio	32.3	36.7	121.1	123.4	510.6	543.4			6.4%	
Oregon	5.1	5.4	47.0	59.5	173.1	200.4			15.8%	
Pennsylvania	40.1	35.4	113.4	126.6	528.1	644.9			22.1%	
Texas	<u>91.8</u>	<u>91.8</u>	<u>344.7</u>	<u>471.1</u>	<u>1,166.7</u>	<u>1,492.8</u>			<u>27.9%</u>	
Non-RGGI Total	201.8	194.0	752.1	901.1	3,049.5	3,567.5			17.0%	
U.S. Total	784.3	737.3	2,936.2	3,090.2	14,798.4	16,094.5			8.8%	

SOURCE: U.S. Bureau of Economic Analysis, Interactive Tables.

TABLE 3
INDUSTRIAL ELECTRIC DEMAND
(Millions of Megawatt-Hours)

	2007	2015	Difference	% Change
RGGI States	52.4	43.1	9.3	-18%
Non-RGGI States	274.6	264.2	10.4	-4%
U.S.	1,027.8	986.5	41.3	-4%

SOURCE: U.S. Energy Information Agency Detailed State Electricity Data.

revenue: \$55 million remains unspent and another \$22 million has gone to administrative overhead and fuel assistance, with just \$23 million (23 percent) going for energy efficiency projects.²

The Maryland Energy Administration (2016) reported that only 25 percent of RGGI revenue was allocated to grants for energy efficiency projects, and that doesn't take into account any money from the grants used for administration by the grantees.

Could the energy efficiency and renewable energy projects have been completed without the RGGI grants? The Maryland 2016

TABLE 4
ENERGY INTENSITY, 2007-15

	Electric Demand Millions MWh		Real GDP \$ Billions		Energy Intensity MWh/\$ Million		
	2007	2015	2007	2015	2007	2015	% Change
RGGI	353.7	342.4	2,338.8	2,504.7	151	137	-9.6%
Non-RGGI	851.9	873.8	3,049.5	3,535.9	279	247	-11.5%
U.S.	3,764.6	3,759.0	14,798.4	16,089.0	254	234	-8.2%

SOURCE: Author calculation dividing electric demand by real GDP.

²Calculation is based on information provided in an unpublished e-mail to a state senator of how DNREC spent RGGI allowance funds from 2014 to 2016, and from SEU Annual Reports (available at www.energizedelaware.org/sustainable-energy).

report, Appendix B, lists hundreds of projects receiving grants. Most of the renewable energy grants went to individuals or companies to install solar photovoltaic cells. The grants were small, running from \$700 to \$1,000 for residential systems that typically cost about \$20,000. Solar projects receive federal tax credits, and the owners can sell renewable energy production credits to utilities that are required to buy them by state law, and receive full credit for every kilowatt-hour of energy produced from the local utility. Using a proprietary spreadsheet program, I find that the internal rate of return of a residential system falls from 10.6 percent with the state grant to 9.2 percent without the grant.³ Most of the projects would move forward without the RGGI revenue grants.

In a report for the Delaware Department of Natural Resources and Environmental Control, Small (2012: 3) found that the federally financed “Weatherization Assistance Program,” which receives 10 percent of RGGI revenue, was shut down for two years while all existing projects were reviewed and redone as needed after a federal audit found various quality control issues. This shows how state evaluation, measurement, and verification measures are not working.

The most rigorous test for energy efficiency projects is to check weather-adjusted meter readings before and after the project is implemented. I have found only one large-scale study by Alberini, Gans, and Towe (2013) that did this. The authors found Maryland homeowners who replaced their heat pumps with no incentives saved an average of 16 percent on electric usage. Meanwhile, homeowners receiving cash incentives of \$300, \$450, and \$1,000 or more had energy savings of 6.2, 5.5, and 0 percent, respectively. The authors concluded on page 7 that “the survey responses provide suggestive evidence the ‘rebaters’ were disproportionately replacing ‘inadequate’ units, leading us to conjecture that the rebates are being used to defray the cost of more powerful units, or of units that end up being used more.”

Table 5 shows predicted changes in electricity demand in the RGGI states based on the 2007 demand adjusted for economic growth (7.2 percent from Table 4), population change (1 percent from U.S. Census data), loss of goods production (−12 percent from

³I assume a 7,500 watt system @\$2.85/watt cost with 20 year life, 9,000 KWh first-year generation reduced 0.5 percent per year, \$0.1425/KWh electric rate rising 2 percent per year, \$6 SREC value, and 30 percent federal investment tax credit.

TABLE 5
PREDICTED CHANGES IN RGGI STATE DEMAND, 2007–15

Cause	%	Change in Demand, millions MWh
Economic Growth	+7.2%	+25
Population Growth	+1%	+4
Loss of Goods- Producing Industry	-12%	-9
Overall Energy Intensity Improvement	-9.6%	-34
Net Theoretical Change		-14
Actual Change		-11

SOURCE: Author calculation multiplying the 2007 demand of 353.7 million MWh times the percentage change in the table, except lost goods production which comes directly from Table 3.

Table 2), and efficiency improvements (-9.6 percent from Table 4). The actual demand fell 11 million megawatt-hours, close to the projected 14 million.

Impact on Carbon Dioxide Emissions

Emissions were reduced about 40 percent from 2007 to 2015 from electric generating units in the RGGI states (Table 6). That compares to only about a 20 percent reduction in emissions for the country as a whole and the comparison states, suggesting RGGI has been a

TABLE 6
CO₂ EMISSIONS FROM POWER PLANTS,
2007–15 (Metric Tons)

	2007	2015	Reduction	% Reduction
RGGI States	144,273,724	87,100,464	57,173,260	39.6
Non-RGGI States	635,998,529	511,342,562	124,655,917	19.6
US Total	2,547,032,486	2,031,452,263	515,558,023	20.2

SOURCE: U.S. EIA (2016a) Detailed State Electricity Data, 1990 to 2015.

TABLE 7
GENERATION MIX PERCENTAGE CHANGE, 2007–15

Fuel	RGGI 2007	RGGI 2015	Non-RGGI 2007	Non-RGGI 2015
Coal	23	7	48	32
Petroleum	3	1	0	0
Natural Gas	32	42	24	33
Nuclear	29	31	22	22
Hydro	10	12	4	3
Other	1	1	1	1
Wind and Solar	0	3	1	6
Biomass and Wood	3	3	1	1

SOURCE: U.S. EIA (2016b) *Electric Power Monthly*.

success. As raw percentages, this would be true, but the base emissions of the RGGI states are much lower than the total for the country, so a relatively small change can appear as a relatively large percent.

Table 7 shows high CO₂ emission coal-fired generation *drops* 16 percentage points in both RGGI and non-RGGI comparison states, and natural gas *rises* virtually the same amount (10 for RGGI states versus 9 for non-RGGI states).

The non-RGGI comparison states actually added more wind and solar generation than the RGGI states (5.5 percentage points versus 2.3), even after allowing for a very large wind farm proliferation in Texas. Some RGGI auction revenue was invested in solar energy projects, but the RGGI, Inc. (2016) report identifies less than 100 MW of added solar capacity, which would account for only about 1 percent of the total wind and solar capacity added in the RGGI states according to generation data in the U.S. EIA *Electric Power Monthly*.

Another way to sort out the impacts of the RGGI program on emissions reductions is to review regulatory and market impacts to the generation mix and emissions in detail. The impacts of exporting emissions through the increased importing of power must also be considered. If a comparison is made of the estimate of emission reductions using just factors common in all states, the comparison

should isolate the impact of the RGGI program. The result of this comparison is discussed below and shows RGGI had no impact on emissions.

Delaware provides an early example of exporting emissions that can be found in a number of articles published in the *Wilmington News Journal* beginning in January 2008. On December 17, 2008, Delaware participated in its first regional cap and trade auction. Three weeks later the Valero-owned Delaware City Refinery announced the shut-down of its electric generation at the plant. According to RGGI, Inc. (2009), CO₂ emissions from the plant's electric generation facility accounted for 17 percent of Delaware's initial emission allocation. Valero had been gasifying petroleum coke, a waste product from the refinery, to fuel the power plant. Petroleum coke has emission rates similar to coal, but by gasifying it Valero reduced emissions of other air pollutants. So, three weeks into the RGGI program Delaware met its total 10 percent RGGI reduction goal. That isn't the end of the story. Valero sold the facility to PBF Energy. PBF restarted portions of the power plant fueled with conventional natural gas. The petroleum coke was loaded onto ships and sent to China to be burned directly for electric generation without pollution controls.

The RGGI states export CO₂ when they increase the import of electricity from other states. Between 2007 and 2015, the RGGI states doubled their imports (Table 8). Much of the imported power comes from the PJM transmission region. Adjusting for this factor decreases the RGGI state emissions reductions about 11 million tons.

TABLE 8
ADJUSTMENT OF RGGI STATE CO₂ EMISSIONS FROM
IMPORTING MORE POWER (Million MWh)

2007 Generation	2007 Demand	2007 Imports	2015 Generation	2015 Demand	2015 Imports	Change in Imports	Added MM tons CO ₂
328.2	353.7	25.5	293	342.4	49.4	23.9	11

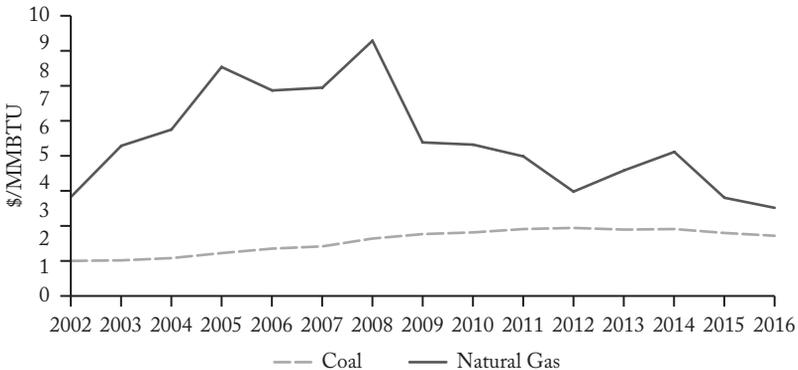
NOTE: Conversion of MWh to metric tons of CO₂ is PJM/EIS (2016) average emission rate of generation of 1,014 pounds/MWh in 2015 divided by 2,204.6 pounds/metric tons or 0.46.

SOURCE: U.S. EIA (2016a) Detailed State Electricity Data, 1990 to 2015.

CO₂ emissions are down across the country. A number of major EPA regulations have been implemented since 2009. Electric power plants have seen the most impact from regulation including the Mercury & Air Toxics Standard (MATS), the Cross State Air Pollution Rule (CSAPR), the Carbon Pollution Standard for New Power Plants that established New Source Performance Standards (NSPS), and the Clean Power Plan (CPP), all aimed at reducing the use of coal and forcing the closure of older, smaller power plants that were not worth upgrading with expensive new filtration equipment, given the low cost of natural gas.

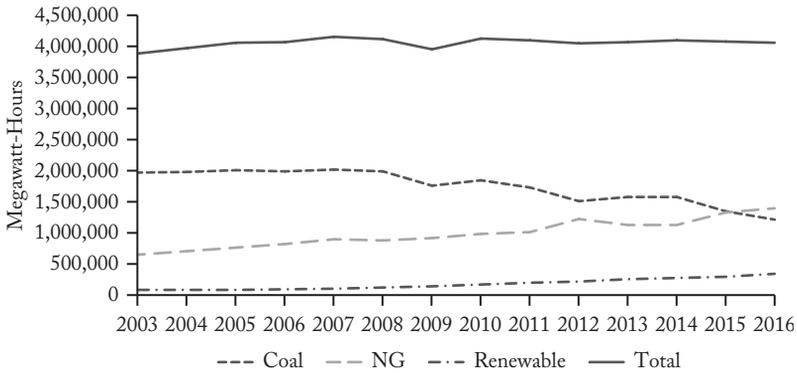
The question is how much of the improvement in power plant emission reduction was caused by EPA regulations. As shown in Figure 3, nominal natural gas prices dropped significantly starting about 2009, driven by an increase in supply from the deployment of hydraulic fracturing and horizontal well drilling technology in shale formations. The types of coal used for electric generation have no other significant uses, and price tends to be stable because electric demand does not vary much from year to year. Natural gas has a number of high volume uses, such as for industrial feedstock and as a primary fuel for heating. Heating demand can vary significantly from year to year. For example, very cold temperatures in the winter of 2014 caused a spike in demand and price. Lower overall natural gas prices played a major role in the switch from coal to natural gas

FIGURE 3
NOMINAL FUEL PRICES FOR ELECTRIC GENERATION,
2002–16



SOURCE: US EIA (2016b) *Electric Power Monthly*.

FIGURE 4
ELECTRIC GENERATION BY FUEL, 2003–16



SOURCE: US EIA (2016b) *Electric Power Monthly*.

for electric generation starting in 2009, and regulations impacted generation capacity starting in 2012.

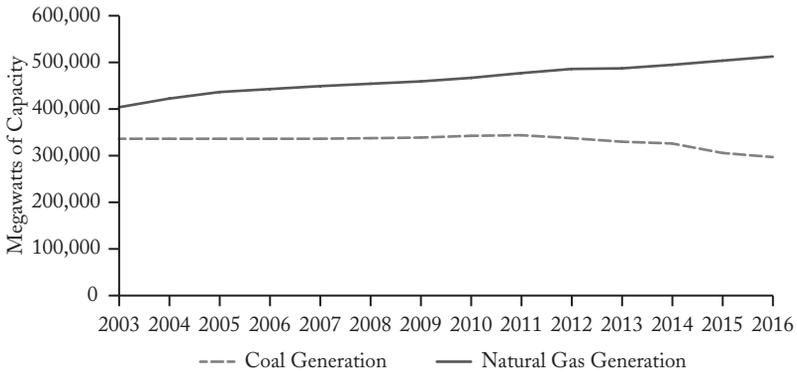
Total electric generation was relatively constant since 2003, but increased almost 3 percent from 2009 to 2016 as the economy recovered from the recession (Figure 4). That increase in demand was met with wind and solar power growth driven by state Renewable Portfolio Standards along with federal and state subsidies. Coal-fired generation was relatively constant until 2008, but began to fall in 2009. The fall paralleled declining natural gas prices. Natural gas generation has been increasing at a relatively constant rate.

EPA regulations did impact coal-fired generation capacity as shown in Figure 5. The downturn in coal capacity coincides with new regulation implementation beginning in 2012. Lower natural gas prices obviously influenced the decisions to close down the coal-fired generation.

However, more important to coal-fired generation was the change in the capacity factor, that is, how often power plants ran in comparison to natural gas-fired power plants (Figure 6). The decline tracks the falling natural gas price curve that began in 2009.

With some certainty nationally, coal plant capacity reductions were caused by EPA regulations, and output reductions were caused by falling nominal natural gas prices. The impact of the two trends can be parsed. The computational details are provided in Stevenson (2017: 12). The result, both nationally and for the RGGI states, is an

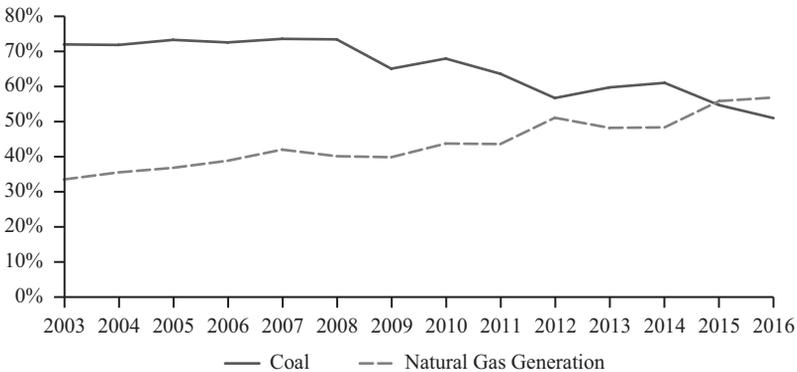
FIGURE 5
ELECTRIC GENERATION BY YEAR, 2003–16



SOURCE: US EIA (2016b) *Electric Power Monthly*.

identical 28 percent from lost generation capacity, and 72 percent from lower natural gas prices. If the RGGI allowance program had a significant impact, it would have offset some of the impact of lower natural gas prices, because the allowance cost acts as an additional variable production cost, and would have shifted the ratio, but it didn't. This result is not unexpected as RGGI allowance revenue only averaged 0.6 percent of electric revenue between 2007 and 2015 (\$0.3 billion/\$51 billion).

FIGURE 6
GENERATION CAPACITY FACTORS, 2003–16



SOURCE: U.S. EIA (2016b) *Electric Power Monthly*.

TABLE 9
CO₂ CALCULATED EMISSION CHANGE COMPARED TO
ACTUAL, 2007–15

Fuel	Est. Lost Generation Millions Megawatt-Hours	CO ₂ Metric Tons/MWh	CO ₂ Emission Reduction MM Metric Tons
Coal	56.0	0.928	-52
Natural Gas	18.2	0.439	+8
Petroleum	5.2	0.901	-4.7
Imported Power	23.8	0.46	<u>-11.0</u>
Calculated Reduction			59.7
Actual Reduction			57.2

SOURCE: Lost generation from U.S. EIA (2016b) *Electric Power Monthly*; emission rates from PJM/EIS (2016).

To complete the estimate of emissions from common factors, the changes in natural gas-fired and petroleum-fired generation need to be added. Table 9 shows that the total net estimated reduction in emissions for RGGI states, due to factors common to all states, was 59.7 million metric tons. That figure is slightly higher than the actual reduction of 57.2 million metric tons, which suggests that the actual reduction is accounted for without any significant additional contribution from the RGGI program.

Low Income Program

According to RGGI, Inc. (2016), in its report titled *The Investment of RGGI Proceeds through 2014*, 15 percent of RGGI revenue (\$178.2 million) went to direct low-income electric bill assistance to 2.6 million households from the beginning of the RGGI auctions in 2008 through 2014. The RGGI funds, about \$30 million a year, were added to the federal Low Income Home Energy Assistance Program (LIHEAP). According to the U.S. Department of Health and Human Services (2014: 10–11), the federal government provided \$795 million to RGGI states in 2014. Thus, RGGI added less than 4 percent to LIHEAP (\$30 million annual RGGI contribution/\$795 million federal contribution).

RGGI allowance revenue totaled \$1.8 billion through 2014. The allowance program added \$0.85/megawatt-hour to electric bills between 2008 and 2014 (\$294 million a year/348 million megawatt-hours demand a year). RGGI state residential electric demand has been fairly flat, and averaged 130.9 million megawatt-hours/year. According to the U.S. Census Bureau (2010), there were 17.3 million households in the RGGI states. Thus, residential electric demand averaged 7.6 megawatt-hours per year ($130.9/17.3$). The total cost of RGGI equaled \$6.50/household ($\0.85×7.6). This reduces the net contribution to low-income households to \$5/year ($\$11.50 - \6.50). Therefore, the net RGGI contribution to the federal LIHEAP was only 1.6 percent, an insignificant amount.

Conclusion

In this article, I investigate claims by the Acadia Center (Stutt, Shattuck, and Kumar 2015: 6) and RGGI, Inc. (2016) that the RGGI program has generated significant benefits. Using data from five comparison states with similar overall electricity policies, except for RGGI, along with looking at national trends, I find the RGGI, Inc. and Acadia Center claims to be misleading.

The Acadia Center claims that compared to other states RGGI states increased electric prices by half as much, had 3.6 percent more economic growth, and reduced emissions 16 percent more leading to greater health benefits from pollution reduction. In reality, from 2007 to 2015, net weighted average nominal electricity prices rose 4.6 percent in RGGI states compared to 2.8 percent in comparison states. Linking real economic growth to RGGI alone is fraught with problems. Real economic growth rates in RGGI states between 2007 and 2015 varied widely from a negative 7.1 percent for Connecticut to a plus 11.9 percent for Massachusetts. Also average RGGI revenue amounted to only 0.01 percent of the combined average real GDP of the RGGI states, so one wouldn't expect much impact. Ignoring those difficulties, real economic growth was 2.4 times faster in comparison states than in the RGGI states. High RGGI state electric rates led to a 34 percent reduction in energy-intensive industries and a 12 percent drop in the goods production sector, while comparison states saw only a 5 percent drop in energy-intensive industries and a 20 percent gain in goods production.

This article finds there were no added reductions in CO₂ emissions, or associated health benefits, from the RGGI program. RGGI emission reductions are consistent with national trend changes caused by new EPA power plant regulations and lower natural gas prices. The comparison requires adjusting for increases in the amount of power imported by the RGGI states, reduced economic growth in RGGI states, and loss of energy-intensive industries in the RGGI states from high electric rates.

The RGGI, Inc. report focuses on the impacts of spending the allowance revenue and suggests significant gains in energy efficiency, wind and solar investments, and assistance with low-income energy bills. Noticeably, RGGI, Inc. does not make claims of superior emission reductions or lower power prices. In reality, the spending of the allowance revenue had marginal impacts. All states have shown energy efficiency gains. The RGGI states saw a lower improvement in energy intensity at 9.6 percent compared to 11.5 percent for comparison states, so there appears to be no RGGI-related gain in overall energy efficiency. Wind and solar energy installation was slower in RGGI states, increasing by only 2.3 percentage points, while comparison states grew by 5.5 percentage points, more than twice as fast. RGGI grants for wind and solar power accounted for only about 1 percent of all the wind and solar power added by the RGGI states. The net fuel assistance help for low-income households, 15 percent of all households, added only 1.6 percent to the federal Low Income Home Energy Assistance Program, or less than \$5/year. RGGI had no meaningful impact on lower-income families. Meanwhile, the other 85 percent of households saw an increase in electricity cost of \$6.50/year directly caused by the RGGI allowance cost.

References

- Alberini, A.; Gans, W.; and Towe, C. A. (2013) "Freeriding, Upsizing, and Energy Efficiency Incentives in Maryland Homes." Fondazione Eni Enrico Mattei Working Paper No. 82. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2336044.
- Brown, J. (2013) "RGGI States Propose Lowering Regional CO₂ Emissions Cap 45%, Implementing a More Flexible Cost-Control Mechanism." RGGI, Inc. Available at [www.rggi.org/docs/Press Releases/PR130207_ModelRule.pdf](http://www.rggi.org/docs/Press_Releases/PR130207_ModelRule.pdf).

- European Commission (2016) “CO₂ Time Series 1990–2015 per Region/Country.” *Emission Database for Global Atmospheric Research (EDGAR)*, release version 4.3.2. Joint Research Centre (JRC)/PBL Netherlands Environmental Assessment Agency.
- Hibbard, P. J.; Tierney, S. F.; Okie, A. M.; and Darling, P. G. (2011) *The Economic Impacts of the Regional Greenhouse Gas Initiative on Ten Northeast and Mid-Atlantic States*. Boston: Analysis Group.
- Maryland Energy Administration (2016) *Maryland Strategic Investment Fund Report on Fund Activities FY2015*. Available at http://energy.maryland.gov/Reports/FY15_SEIF_Annual_Report.pdf.
- PJM/EIS (2016) “Generation Attributes Tracking System.” *System Mix 2015*. Available at <https://gats.pjm-eis.com/GATS2/PublicReports/PJMSystemMix/FilterPJM>.
- RGGI, Inc. (2009) “RGGI Historical Emissions 2000 to 2008 Spreadsheet.” Available at http://www.rggi.org/historical_emissions.
- _____ (2016) “The Investment of RGGI Proceeds through 2014.” Available at www.rggi.org/docs/ProceedsReport/RGGI_Proceeds_Report_2014.pdf.
- _____ (2017) “Auction Results for Auctions 1–36.” Available at www.rggi.org/market/co2_auctions/results.
- Shultz, G. P., and Summers, L. (2017) “This Is One Climate Solution That’s Best for the Environment and for Business.” *Washington Post* (June 19).
- Small, D. (2012) *Annual Report on the Weatherization Assistance Program*. Wilmington, Del.: Delaware Department of Natural Resources & Environmental Control.
- Stevenson, D. T. (2017) “Sorting Root Causes of Air Quality Improvements 2009 to 2015.” Caesar Rodney Institute Policy Paper (January). Available at <https://criblog.wordpress.com/2017/02/12/sorting-root-causes-of-air-quality-improvements-2009-to-2015>.
- Stevenson, D. T., and Stapleford, J. E. (2016) “Cost Impacts of 2013 RGGI Rule Changes in Delaware.” Caesar Rodney Institute Policy Paper (August). Available at www.caesarrodney.org/pdfs/Cost_Impacts_of_2013_RGGI_Rule_Changes_in_Delaware.pdf.
- Stutt, J.; Shattuck, P.; and Kumar, V. (2015) *Regional Greenhouse Gas Initiative Status Report, Part 1: Measuring Success*. Boston: Acadia Center (July).

- Tung, M. B. (2017) “RPS Requirement and Aggregated Cost of RPS.” Presentation to the House Economic Matters Committee, Maryland Energy Administration, Baltimore, Md. (January 15). Available at <http://news.maryland.gov/mea/wp-content/uploads/sites/15/2017/01/EconomicMattersPresentation1.23.2017.pdf>.
- U.S. Bureau of Economic Analysis (n.d.) “Real Chained GDP by State.” Interactive Tables available at www.bea.gov.
- U.S. Census Bureau (2010) “Population and Housing Unit Estimate Tables by Year.” Available at www.census.gov/programs-surveys/popest/data/tables.2010.html.
- U.S. Department of Health and Human Services (2014) *LIHEAP Report to Congress for Fiscal Year 2014*. Division of Energy Services. Washington: DHHS.
- U.S. Energy Information Agency (EIA) (2016a) Detailed State Electricity Data: “Average Price by State by Provider (EIA-861)”; “Revenue from Retail Sales of Electricity by State by Sector by Provider (EIA-861)”; “Retail Sales of Electricity by State by Sector by Provider (EIA-861)”; “Net Generation by State by Type of Producer by Energy Source (EIA-906, EIA-920, and EIA-923)”; “U.S. Electric Power Industry Estimated Emissions by State (EIA-767, EIA-906, EIA-920, and EIA-923).” Available at www.eia.gov/electricity/data/state.
-
- _____ (2016b) *Electric Power Monthly* (2007 and 2015): “Table 1.4.B Coal by State by Sector, Year-to-Date”; “Table 1.5.B Petroleum Liquids by State by Sector, Year-to-Date”; “Table 1.7.B Natural Gas by State by Sector, Year-to-Date”; “Table 1.9.B Nuclear Energy by State by Sector, Year-to-Date”; “Table 1.10B Hydroelectric (Conventional) Power by State by Sector, Year-to-Date”; “Table 1.14.B Wind by State by Sector, Year-to-Date”; “Table 1.17.B Solar Photovoltaic by State by Sector, Year-to-Date”; Table 4.2 Total Average Cost of Fossil Fuels for Electric Utilities”; “Table 6.7A Capacity Factors for Utility Scale Generators Using Fossil Fuels.” Available at www.eia.gov/electricity/monthly.

A TEACHER'S PERSPECTIVE ON WHAT'S WRONG WITH OUR SCHOOLS

Peter McAllister

The American public has been aware of the poor performance of our schools since the Reagan Administration published *A Nation at Risk: The Imperative for Education Reform* in 1983. Since then, educational reformers have engaged in several major programs, such as the No Child Left Behind Act and the Common Core State Standards Initiative. Despite these efforts, the performance of our schools is still unsatisfactory. We have one of the highest levels of educational spending per student in the world, yet American students' performance in international comparisons of educational quality is consistently mediocre. According to the National Assessment of Educational Progress (2017), only 25 percent of American 12th grade students are proficient in math, 22 percent are proficient in science, and 12 percent are proficient in U. S. history.

Members of several professional groups, including economists, psychologists, teacher union representatives, public policy analysts, and politicians, have opined upon our educational policies for decades. Teachers, on the other hand, have been virtually silent. Yet teachers have direct experience of the practical consequences of our educational policies, and are in a position to identify problems in our

Cato Journal, Vol. 38, No. 1 (Winter 2018). Copyright © Cato Institute. All rights reserved.

Peter McAllister taught mathematics and social studies at West Islip High School in Long Island, New York. Before becoming a teacher, he was a Vice President of Citibank. He was also a cofounder of the American Renaissance School, a for-profit high school that was located in White Plains, New York. The author would like to thank Ronald Pisaturo, James Bencivenga, Donna Wilder, and Donald Dutkowsky for their insights during the preparation of this article.

schools that may not have been discussed adequately in the public forum.

Based upon my own teaching experience, I believe that the discouraging test performance of American students is a symptom of fundamental problems in our schools. When I taught mathematics and social studies in a public high school, I observed the following characteristics in many of my students:

- *A nonconceptual mentality.* Students could not integrate the information that they learned into a comprehensive understanding of a subject. Because they could not comprehend their subject matter, they simply memorized it.
- *A lack of independent thought.* When drawing conclusions, students rarely relied on their own informed judgment. Instead, they either relied on their emotional reactions or uncritically adopted the opinion of the group.
- *Ethical expediency.* Students frequently cut corners for the sake of some short-term gain. If they thought that the rewards were high enough, students violated rules that they would not have violated otherwise.

In this article, I argue that the policies and practices that engendered these characteristics in my students are based upon the pedagogical principles of “progressive education”—that is, the educational philosophy that has dominated mainstream American education for the past century. Although progressive education ended as a formal movement in the 1950s (Cremin 1961: 347–53), state departments of education have ensured that progressive educational principles have remained firmly entrenched within our schools. These principles, applied faithfully by American educators for generations, have produced students who do not think, who do not question, and who do not care about what is right or wrong.

Nonconceptual Mentality

Students have difficulty learning to think conceptually because many of our curricula violate a fundamental principle of epistemology: the hierarchical structure of knowledge.

In her *Introduction to Objectivist Epistemology*, Ayn Rand (1966) identified a crucial aspect of conceptual knowledge: concepts have a hierarchical structure. Higher-level concepts are derived from

simpler, more basic concepts. For example, it is impossible for a child to grasp the concept of “furniture” without first grasping more basic concepts from which the concept furniture is derived such as “bed,” “table,” “chair,” and so on. There is a necessary order in the sequence of abstractions that a child must follow to form this concept. This necessity of order applies to every concept. A higher-level concept, such as “electron,” requires a lengthy, complex sequence of abstractions. Each step in this sequence must be performed in a hierarchical order. If any step is omitted or not performed properly, the resulting idea cannot be considered valid conceptual knowledge.

This aspect of Ayn Rand’s theory of concepts has a clear pedagogical implication: because there is a necessary order in which concepts are formed, there is a necessary order in which concepts must be taught. As Lisa VanDamme (2006: 59), the headmistress of the VanDamme Academy, has pointed out, “An abstract idea—whether a concept, generalization, principle or theory—should never be taught to a child unless he has already grasped those ideas that necessarily precede it in the hierarchy, all the way down to the perceptual level.” Because the order of presentation of conceptual knowledge must follow the principle of hierarchy, a proper education requires a structured curriculum that also follows this principle. If the proper order of presentation of a subject is violated, it is often impossible for a child to integrate new information into his knowledge base.

Progressive educators, however, have rejected structured, hierarchical curricula. William Heard Kilpatrick, who was reputed to have trained 35,000 teachers during his tenure at Teachers College, Columbia University (Ravitch 1983: 50), was outspoken in his antipathy toward any structured curricula. Kilpatrick (1925: 266) advocated an education that “would stress thinking and methods of attack and principles of action” rather than curricula that he derisively referred to as “subject matter fixed-in-advance.” This rejection of fixed curricula led to the separation of thinking methods from content, with content relegated to a secondary status. Under the influence of Kilpatrick and other progressive educators, teachers gradually abandoned courses with content presented in a logical hierarchy.

The U.S. experience with mathematics education is an obvious example of this development. State departments of education began to mandate instruction in “New Math” in the late 1950s and, by the late 1960s, the New Math curriculum had become prevalent in our

schools. As Morris Klein (1973) has pointed out, this curriculum repeatedly violated the principle of hierarchy. Elementary school students struggling to grasp the rudiments of arithmetic have had to learn aspects of number theory that are meaningless to children of their age. For example, in many schools, instead of teaching subtraction by simply counting and removing quantities of real objects, such as tiles or beads, elementary school teachers taught their students that the operation of subtraction entails the addition of a *negative quantity*. It took the world's leading mathematical thinkers a millennium to develop the concept of a negative quantity, and another millennium to accept it (Klein 1973: 40); yet mathematics teachers have expected children somehow to master this counterintuitive concept in a matter of weeks. Similarly, even though it took 22 centuries for mathematicians to notice that Euclid's *Elements* rested on more fundamental axioms than Euclid stated, geometry teachers have expected teenagers not only to grasp this shortcoming, but also to understand why they must use abstruse axioms to prove the simplest of theorems. New Math curricula have flitted across exotic, unrelated topics such as set theory, locus, symbolic logic, and linear algebra without providing students with any indication of the practical application of these topics or a logical context within which to integrate them.

To get an idea of how esoteric New Math curricula have been, consider the topic of material implication. This topic deals with conditional sentences, which are statements in the form of "if . . . , then." Material implication obviates any causal relationship in these statements. Material implication holds that the only time a conditional sentence is false is when the antecedent (the "if" part of the sentence) is true and the consequent (the "then" part) is false. All other conditional sentences are true. Thus, a sentence such as "If Paris is the capital of France, then the sun rises in the West" is false, not because it makes no sense, but because the antecedent is true (Paris *is* the capital of France) and the consequent is false (the sun *does not* rise in the West). On the other hand, the sentence "If the sun rises in the West, then Paris is the capital of France" is a true statement. Try explaining that to a class of ninth graders! Material implication first gained notoriety when Bertrand Russell included it in *Principia Mathematica* (Whitehead and Russell 1910: 98). Since then, philosophers have challenged the utility of this logical formalism (Blanshard 1939: 374–81). As you may imagine, most well-educated adults

would have a great deal of difficulty understanding this topic. You also may be wondering what any of this has to do with mathematics. Yet material implication was part of New York State's Integrated Mathematics curriculum for decades and appeared in widely used mathematics textbooks (Dressler and Keenan 1989: 147–51). The New York State Regents examinations regularly included questions involving material implication, or topics derived from it, until January 2009. It has also been part of the curricula of school districts in California and several other states. Material implication is not part of the Common Core standards, however, and it is no longer included in the mathematics curricula of either New York or California.

By discarding the hierarchical, logical chain that connects mathematical abstractions to physical reality, American educators have severed mathematics from the real world in the minds of many of their students. I encountered the consequences of this separation shortly after I began teaching high school mathematics. When I tried to show my students how the topics we were covering have real world, perceptual foundations, students typically demanded that I stop and just give them a formula that they could memorize. They had no interest in grasping the hierarchical chain that connects mathematical principles to observable facts. Based upon their prior mathematics courses, my students had come to expect instruction in mathematics to consist of manipulating a series of strange little symbols whose meaning defy comprehension. I also discovered that, because mathematics teachers have learned this subject the same way they teach it, many mathematics teachers do not understand the real world foundation of mathematics either.

Consequently, mathematics education in our schools has often devolved into a process of cramming formulas into students' minds, with mnemonic devices substituting for instruction. For example, instead of teaching the derivation of the quadratic formula—a formula indispensable in the design of objects ranging from flashlights to the Golden Gate Bridge—teachers often “teach” this formula by singing it with their students to the tune of *Pop Goes the Weasel*. Mathematics teachers have relied on similar approaches to teach topics in number theory, algebra, geometry, and trigonometry.

Before I became a teacher, I spent more than two decades setting up and managing departments that specialized in using applied mathematics to solve business problems for a multinational bank. During that period, I often had difficulty finding Americans who

possessed the mathematical skills that I required. Shortly after becoming a teacher, I learned why: instruction in our state-mandated mathematics curricula has caused this subject to become unfathomable to many young minds.

The opposition to hierarchical curricula on the part of progressive educators has permeated instruction in other subjects as well. A proper history curriculum should begin with objective facts and proceed in a chronological sequence, so that students can see how choices made by people in one period cause outcomes in the next. When education historian Diane Ravitch (1985) visited a local New York City high school in the mid-1980s, she was surprised by the complete absence of traditional, chronological instruction in history. The department chairman informed her that, since the late 1960s, the school had followed the guidelines of the New York State Education Department, which was to deemphasize chronological history and to focus instead on topical issues and social science concepts.

A teacher in the same school pointed out that he did not teach history because it did not help his students pass the New York State Regents examination in social studies. Under the influence of progressive educators, such as John Dewey and Harold Rugg, school districts have eliminated history departments and replaced them with social studies departments. In addition to teaching courses in history, many social studies departments now offer courses in economics, sociology, anthropology, and psychology. Ravitch noted that the introduction of these other disciplines caused considerable confusion among teachers and administrators as to what constitutes a proper curriculum; it also left little time for hierarchical, chronological instruction in history.

Like mathematics courses, science courses should begin with perceptually verifiable information before proceeding to higher levels of abstraction. Instead, for decades, science teachers have plunged students into worlds of subatomic particles, obscure biological processes, and ineffable forces with little attempt to show how such worlds were discovered from more basic knowledge or how the information that students are expected to learn has practical, observable consequences. VanDamme (2006) has pointed out that this type of instruction in science paradoxically requires a nonscientific act of faith on the part of the student. For example, when a child is presented with a schematic representation of an atom, with hardly any

hierarchical explanation of the process by which atomic theory was developed, the child has no more compelling reason for believing in atomic theory than he has for believing in Santa Claus.

State departments of education also have virtually eliminated instruction in grammar from the public school curriculum (Mulroy 2003), and have introduced the nonphonetic—that is, nonconceptual—whole word approach to reading instruction as either an alternative or a supplement to phonics-based instruction in reading.

Ironically, by denigrating hierarchical course content, American educators have made students far less able to master the thinking methods prized by progressives. The curriculum, the very content upon which students might learn to apply essential methods of thinking, militates against the development of these methods. With method divorced from content, students have come to view thinking methods as arbitrary rules of an unreal game. Instead of trying to understand these rules, students simply memorize them—exactly the opposite of the result sought by Kilpatrick and his progressive colleagues.

This substitution of memorization for understanding has inculcated the nonconceptual mentality in many students. These students can often retain large amounts of information and perform calculations and other routine mental operations rapidly, but they can do very little conceptualizing. They memorize, manipulate, and recall information well; they just cannot think. Because they do not grasp the difference between understanding and memorization, they are often incapable of identifying when they do not know something.

Although educators have been aware of the need of a hierarchical approach to curriculum development, based upon my experience as a teacher and my review of the relevant literature, I believe that they have not fully realized its critical importance. In 1956, psychologist Benjamin Bloom (1956) chaired a committee of educators who developed a set of hierarchical models that classify learning objectives into levels of complexity and specificity. Robert M. Gagne (1965) developed a similar system of analyzing learning from simple to complex levels. These taxonomies present general guidelines regarding the order of presentation of knowledge, and they have been used extensively in the design of curricula. There is also a substantial amount of material that discusses, in general, the relationship of curricula to students' acquisition of conceptual knowledge. However, few researchers have sought to bridge the gap between

the hierarchical structure of concepts within epistemology and the development of a hierarchical curriculum within pedagogy. The only curricula that I am aware of that present content in a comprehensive, hierarchical manner are the Core Knowledge Sequence developed by E.D. Hirsch and the Core Knowledge Foundation, and the curricula in place at the VanDamme Academy.

There is another aspect of our public education system that militates against the establishment of a hierarchical approach to education: most public school teachers are not experts in subject matter. Beginning in the early 20th century, Ravitch (2003) notes, educators sought to establish teaching as a distinct profession, similar to law or medicine. Teacher certification became dependent upon taking courses in pedagogy and passing exams in pedagogic theory, and less dependent upon mastery of the subject taught. A report produced by the National Center for Educational Statistics (Lewis et al. 1999: 12–13) found that, by the end of the 20th century, only 38 percent of teachers held a bachelor's or graduate degree in an academic field other than education. The report noted that this statistic was unchanged from a similar survey conducted five years earlier. A more recent report by the U. S. Department of Education (Staklis and Henke 2013: 9) showed that one half of the 2007–08 bachelor's degree recipients who taught the following year had majored in education. To inculcate a hierarchical understanding of the subjects they teach, teachers themselves need a comprehensive, hierarchical understanding of these subjects. To acquire this understanding, teachers need, at least, to have attained a bachelor's degree in these subjects.

On a topic related to hierarchy, E.D. Hirsch (2016: 13) notes that current research disputes progressive educators' justification for what he refers to as "skills-centrism": the belief that teachers can impart general thinking skills, such as critical thinking and problem solving, without regard to specific subject content. He cites the comprehensive *Cambridge Handbook of Expertise and Expert Performance* (Feltovich, Prietula, and Ericsson 2006: 11), which states, "Research clearly rejects the classical views on human cognition in which general abilities such as learning, reasoning, problem solving, and concept formation correspond to capacities and abilities that can be studied independently of the content domain." The decoupling of thinking skills and subject matter that led to the proliferation of the nonhierarchical curricula discussed above rests upon this discredited assumption of skills-centrism.

Hirsch (2016) also presents a cogent argument for the failure of progressive educational ideas in his discussions of the decline in American schooling in the 1960s and 1970s, and the decline in French and Swedish schooling underway currently. Hirsch maintains that even though progressive ideas had begun to take over American public schools in the 1920s, it took decades for these ideas to achieve dominance. Although the progressive education movement ended in the 1950s, Hirsch notes, by the 1960s progressive ideas regarding teaching methods and curricula had become conventional educational practice in most American schools. Between 1960 and 1980, American academic scores fell rapidly. Academic scores across all grade levels fell more than a quarter of a standard deviation, a big decline for large populations. SAT scores fell by half a standard deviation. Student performance on other national tests such as the ACT (a test similar to the SAT), the Iowa Test of Educational Development, the Minnesota College Aptitude Test, the California Achievement Test, and the Stanford Achievement Test all showed large declines (Hirsch 2016: 133)

Hirsch describes the same pattern emerging in France and Sweden. In 1989, after decades of growing interest in progressive educational ideas in France's schools of education, France adopted the *loi Jospin* ("the Jospin Law"); this law effectively required that France's elementary schools institute a child-focused, skills-centric approach to education similar to the approach in American schools. A sample of the performance of 300,000 French fifth grade students in 2007 showed more than a third of a standard deviation decline in reading skills and more than four-fifths of a standard deviation decline in mathematics skills relative to the performance of students in 1987—two very large, very significant declines. When stratified by demographic characteristics, the results showed declines for all groups, with the largest declines in those groups within the lowest socioeconomic statuses (Hirsch 2016: 211). In 1994, Sweden made changes to its educational system similar to those made in France. From 2000 to 2012, Sweden's scores on the PISA test, a frequently cited assessment of national educational quality, dropped from 516 to 483 in reading, from 510 to 478 in mathematics, and from 512 to 485 in science. These declines were monotonic: the result of each test in each category was lower than the previous result (Hirsch 2016: 154). Apparently, France and Sweden are determined to duplicate America's educational mistakes.

It is our children who suffer because of these ideas. In a recent multiyear college and career readiness survey of 165,000 high school students, 87 percent of the respondents stated that they hope to earn a college degree and embark upon a career, but only 45 percent feel positively about their college and career readiness (YouthTruth 2015). A separate survey, consisting of interviews with college instructors and employers (Achieve 2015), indicates that these students have reason to be concerned. The survey found that 78 percent of college faculty and 62 percent of employers believe that public high schools are not preparing students for the demands that these students will face when they graduate.

Lack of Independent Thought

One of the salient characteristics of progressive education is a shift away from lecture-based, teacher-centered instruction to group-based, student-centered instruction. Based on my experience, this change in focus, when combined with the effects of nonhierarchical curricula, has made students reluctant to question the content that is presented to them or to investigate issues thoroughly when such an investigation may cause them to challenge opinions that they consider part of the status quo.

John Dewey, “the father of progressive education,” believed that guiding a child’s adjustment to the demands of the group is one of the most important responsibilities of an educational system. According to Dewey (1897: 77), “The only true education comes through the stimulation of the child’s powers by the demands of the social situations in which he finds himself. Through these demands he is stimulated to act as a member of a unity . . . to conceive of himself from the standpoint of the welfare of the group to which he belongs.”

Under the influence of Dewey and other progressive educators, educational administrators began requiring teachers to focus on conforming their students’ behavior to group norms during the early 20th century. This development was most pronounced within the elementary grades, where students’ social interactions are most pliant. Dewey replaced permanently anchored classroom desks in his Laboratory School in Chicago with movable units that could be rearranged at will to suit the needs of the group (Rugg and Shumaker 1928: 1). School districts across the country followed suit.

In leading progressive schools, the focus of instruction shifted from that of students acquiring individual intellectual skills to students completing group projects. In the Lincoln School in New York City, elementary school students decided what projects the group would work on and who would perform what tasks. The teacher was only “a wise but inconspicuous member of the group” (Rugg and Shumaker 1928: 57). By the end of the 20th century, group work had become common within all grades in our schools. A report on teacher development by the National Center for Educational Statistics (Choi and Chen 1998: 45) found that, during the 1993–94 school year, 49 percent of teachers reported that they had received training in methods of cooperative learning wherein students work together in groups to solve a problem or produce a product.

As a high school teacher, I was frequently directed to avoid content-based lecturing in favor of group activities. In my education courses, particularly those courses concerned with classroom management skills, my professors stressed the advantages of allowing groups of students to work independently on projects (Cooper 1999: 267–307). As part of my state certification requirements, I had to submit videotaped lessons to the New York State Department of Education. The teacher who supervised the production of these videotapes informed me that, unless at least one-third of each lesson contained group activities, my application would likely be rejected. In my performance reviews, school administrators told me to minimize time spent lecturing in favor of allowing students to work independently in groups. Administrators urged me to be “the guide on the side, not the sage on the stage.” In department meetings, my colleagues frequently praised a variation of the Socratic method they used known as Socratic Circles, which effectively allow students to control the direction and the flow of knowledge in the classroom.

Despite rosy descriptions of shifting the focus of instruction from the teacher to the student, I found that this group approach to instruction yielded discouraging results. Many students became free riders in these groups and contributed nothing, causing those students who worked on the group assignment to become resentful. I had to spend a great deal of time keeping students on task and preventing these group activities from becoming social get-togethers. In general, the quality of student work was poor. When students work together in groups, they often get bogged down with superficial issues and rarely think in fundamental terms. Used judiciously,

carefully guided group work can be effective; but, overall, students cannot direct the course of their own instruction.

This type of instruction when combined with the effects of non-hierarchical curricula instilled conformity in many of my students. Students frequently saw no need to justify their opinions on any basis other than their emotions, or they uncritically accepted the group consensus. This behavior is understandable. Many students, for example, do not have a chronological, hierarchical grasp of historical facts and yet are asked to develop recommendations regarding issues such as income inequality, nuclear nonproliferation, and the Arab-Israeli conflict. How can we expect these students, who have hardly any knowledge of what men and women have actually done, to make well-reasoned recommendations regarding what men and women should do?

Students who do not think independently are defenseless against a subtle, but growing, form of propaganda within our schools. In many introductory economics courses, teachers coach students with no previous knowledge of financial markets or economic principles to replicate complex explanations of how the Federal Reserve's purchase and sale of securities affect interest rates, foreign exchange rates, the levels of consumption and investment, and ultimately nominal and real GDP. These financially nascent students must also provide critiques of rarefied and controversial topics such as the central positions advanced by the Keynesians and Monetarists. Widely used texts for these courses focus on the mechanical aspects of these topics, with only a cursory examination of fundamental economic principles (Clayton 1995, Baumol and Blinder 2005, Carbaugh 2008). By compressing so much sophisticated information into a high school course, these courses frequently violate the principle of hierarchy. Consequently, many students simply memorize this material, and they do not question it. The often unstated premise in these courses and their supporting textbooks is that, without the intervention of the Federal Reserve and the fiscal policies of the federal government, our economy would lurch from crisis to crisis. Because many students equate memorization with comprehension, they accept this premise as true. Similarly, students learn about problems caused by monopolies, oligopolies, and monopsonies, with the unstated premise that these market structures are the unavoidable consequences of an unregulated economy. Students accept this premise as well. Social studies departments portray these courses as unbiased, on the

grounds that the textbooks and syllabi of these courses merely reflect the economic reality of our mixed economy. Having taught these courses for several years, I can testify that the consistent message students receive is that, without the extensive and often drastic intervention of government policymakers and regulators, our economy would collapse.

Students receive similar messages in their other social studies courses. The authors of American history textbooks, from texts used in the 1960s (Planer and Neff 1962) to texts used currently (Brinkley 2007), have portrayed the late 19th and early 20th centuries as a period of escalating class conflict during which politicians had to protect the working classes from the predatory behavior of industrialists. Historian Larry Schweikart (2008) examined several of the most widely used American history textbooks in upper-level high school and introductory college courses and found evidence of an increasing bias in favor of a progressive, left-liberal interpretation of our past. He found that these texts maintain that the Great Depression was caused by the tax cuts and unfettered flow of capital of the late 1920s, that FDR's New Deal and LBJ's Great Society were major advances in American society, and that Northern capitalist greed—not slavery—drove the civil war. These textbooks also claim that our transcontinental railroads would never have been built without the intervention of the government, and that it was the lack of government intervention that caused the declines in the U.S. automobile and steel industries.

In literature courses, teachers often blame the avarice of Britain's industrialists for the miserable social conditions described by Charles Dickens and other writers, and claim that these conditions ended only when legislators enacted child labor laws and other social legislation. In science courses, teachers claim that federal agencies must work constantly to prevent industrialists from destroying our environment. Since 2007, every student I taught had seen *An Inconvenient Truth* in at least one of his science courses, but few students were aware that there are scientists who dispute the claims made in this documentary. I do not know how many other school districts have shown their students *An Inconvenient Truth*, but Schweikart points out that students are exposed to other sources, including American history textbooks, that imply that man-made global warming is real and caused primarily by insufficiently regulated American industries. Even some mathematics textbooks now exhibit an ideological

point of view. In his textbook *Precalculus*, for example, Robert Blitzer (2001: 362–63) illustrates how exponential functions can be used to calculate how rapidly the world’s population would grow without some type of presumed intervention to prevent the exhaustion of our global resources.

The theme that runs through these lessons, and a myriad of other lessons, is the standard progressive conceit: we cannot trust private individuals to act on their own. Instead, we need government experts to guide our choices, and politicians to protect us from the baleful consequences of capitalism. Isabel Paterson (1943: 258) warned that “every politically controlled educational system will inculcate the doctrine of state supremacy sooner or later . . . Once that doctrine has been accepted, it becomes an almost superhuman task to break the stranglehold of the political power over the life of the citizen.” Over the 75 years since Paterson issued her warning, the movement of our public schools in this direction has been unmistakable.

Ethical Expedience

John Dewey rejected the idea of moral absolutes. According to Dewey (1929: 222), “A moral law, like a law in physics, is not something to swear by and stick to at all hazard.” Instead, to Dewey and his progressive followers, efficacy is the primary criterion of ethical validity. Moral rules should be treated “as intellectual instruments to be tested and confirmed—and altered—through consequences effected by acting upon them” (Dewey 1929: 221). This progressive emphasis on efficacy, as opposed to strict adherence to rules, has been captured in a four word slogan that has echoed throughout our schools for more than half a century: “There are no absolutes.”

It is impossible to make a blanket statement regarding how much the environment of our schools has changed because of Dewey’s emphasis on efficacy. At the beginning of the twentieth century, Dewey (1915: 19) himself estimated that only 1 percent of the American population attended college and only 5 percent entered high school. The dramatic increase in school attendance during the 20th century, along with waves of immigration, exerted profound effects upon the conditions in our schools. Regardless of the sources of these changes, however, our schools have experienced a definite movement away from rigorously enforced codes of conduct and academic standards.

Because rules are absolute—either you make it to class by the bell or you do not—formerly strict, traditional codes of conduct in many schools have become guidelines, open to continuous modification and interpretation. I am personally aware of administrators who, because it would be inefficacious to “rock the boat,” have chosen not to deal with violations of conduct codes properly in situations ranging from insubordination, inappropriate dress, and cell phone usage, to systematic cheating and assault. Additionally, students know that they can reduce or overturn almost any punishment for any infraction of school policy if they can persuade their parents or guardians to lodge a formal protest against it. A 2004 survey (Public Agenda 2004: 4) found that 55 percent of teachers believe that school districts’ backing down when confronted by assertive parents causes discipline problems in their classrooms.

Absolute academic standards, likewise, have often disappeared. Grade inflation has been chronic, and school systems have adopted policies such as social promotion, whereby grade advancement is determined by a student’s age rather than his intellectual development. Because it takes a lot more time to fail a student than to give him a marginally passing grade, and because high failure rates spell trouble for school districts, teachers and administrators sometimes find it expedient to avoid adhering strictly to grading criteria. In states where students are required to pass standardized examinations to advance academically, teachers routinely have been able to find extra points for failing students. This practice became so widespread that the New York State Department of Education had to introduce elaborate grading procedures for its state Regents Examinations so that teachers could no longer change any student’s examination grade once it was computed. School district administrators frequently fail to enforce attendance policies, and teachers have even been ordered by their school board to overlook plagiarism policies (Wilgoren 2002).

Many teachers who embrace multicultural education have communicated moral relativism explicitly in the classroom. Psychologists Blaine Fowers and Frank Richardson (1996: 609) define multiculturalism as “a social-intellectual movement that promotes the value of diversity as a core principle and insists that all cultural groups be treated with respect and as equals.” Teachers who engage in multicultural education often do so because they sincerely believe either that cultures cannot be evaluated according to any absolute, objective

criteria, or that it is wrong to do so. Others, however, may simply find it expedient to retreat to multiculturalism to avoid thorny issues related to criticizing other cultures. Either way, if all cultural groups are to be treated as equals, then the only moral conclusion students are likely to reach in these classes is “you must not judge.”

It is difficult to assess the moral caliber of youth, and even more difficult to ascribe specific moral characteristics to the influence of schooling. Individuals are free to choose their own actions; and other factors, such as family life and religious affiliation, exert powerful influences upon the development of an individual’s character. Despite these disclaimers, it is hard to imagine that students have not been affected by the progressive influence on the moral environment of our schools. Before progressive ideas became influential, many traditional schools emphasized to their students that it is right for students to base their actions on moral principles; the schools backed up this emphasis with strictly enforced codes of conduct. This situation, however, was not what I encountered when I taught. Because teachers and administrators, both explicitly and by their actions, had assured students that “there are no absolutes,” many students regarded issues of right and wrong as mostly irrelevant. If the likelihood of getting caught for a tempting infraction of school policy was high and the penalty was great enough, students did not break the rules. Otherwise, they did. If I left a classroom where I was proctoring an examination, students immediately began to cheat. If calculators, cell phones or other valuable items were left unattended, students were likely to steal them. This unprincipled behavior is consistent with progressive educators’ ethical pragmatism. If there are no moral absolutes, then ethics is a matter of expediency.

Conclusion

In *Educational Wastelands*, published in 1953, Arthur E. Bestor (1953) accused progressive educators of abdicating the responsibility to teach the “power to think.” That same year, Robert M. Hutchins (1953), in *The Conflict in Education in a Democratic Society*, criticized what he referred to as progressive educators’ “doctrine of adjustment,” which prized conformity and devalued independent thought. Albert Lynd (1953: 36), in *Quackery in the Public Schools*, also published in 1953, characterized the methods of teacher training and certification, wherein teachers are taught and certified only by

teachers who are themselves steeped in progressive educational theory, as “one of the neatest bureaucratic machines created by any professional group in any country anywhere.” The specific conditions that these authors criticized have changed but, 65 years later, the underlying problems that they identified continue to plague our educational system.

The foregoing leads to the following question: Is the central problem in our schools the influence of progressive educational ideas, or is it the patchwork of state education monopolies that has rendered our schools impervious to fundamental reform? These well-coordinated monopolies have suppressed the marketplace for new ideas in education, and they have pitted the rent-seeking behavior of teacher unions against the well-being of our children.

I believe that our current system of state education monopolies is the biggest obstacle we face to achieving genuine educational reform. In the Winter 2016 edition of the *Cato Journal*, Benjamin Scafidi (2016: 134–38) discussed a proposal offered by economist Richard Vedder that may overcome some of the resistance to dismantling these monopolies. Vedder (2000) proposed using “Employee Stock Ownership Plans” to convert individual public schools into autonomous, employee-owned enterprises. These schools would have control over the curricula, teaching methods, and conduct codes that they adopt. The tuition charged by these schools would be covered by voucher payments, but Vedder’s proposal does allow for parents to bid up the price of tuition by adding their own funds to the amount of tuition they offer to pay. Although students from low-income families and special needs students may have to receive additional resources, a trial program could be structured to identify and address problems that these groups of students encounter. Because teachers and other stakeholders would become owners of these schools, their resistance to ending state control of education may be reduced.

Vedder’s approach provides the means to address all the issues I have identified above. Most critically, it would end the states’ monopolies of educational policies. If parents still want their children to receive an education based upon progressive principles, they would be free to send their children to a progressive school. In fact, Vedder speculated that most schools would not initially offer programs of instruction that were radically different from the those in place currently. Over time, I believe this would change. Private schools need

the approval of state departments of education to confer diplomas recognized by the state. If the need for this approval was eliminated, I believe that many schools would be willing to explore innovative educational approaches. This could lead to a dramatic improvement in the quality of American schools.

The alternative to trying Vedder's approach, or a similar approach, is to continue the pitched battles that proponents of school choice are engaged in currently with teacher unions and the politicians that these unions support. Unfortunately, these battles will leave our counterproductive educational policies mostly untouched.

When I was a teacher in a government-run school, I witnessed firsthand the consequences of the educational policies discussed in this article. If we wish to realize the schools that our children deserve, we need to jettison these educational policies along with the state monopolies that have preserved them.

Our children deserve schools that train them in the proper development and use of concepts; that facts and principles cannot be evaded with impunity; and that their own unaided reason, rather than any majority's opinion or anyone's emotional reaction, is the only means whereby they can discover the truth. And our children deserve schools that demonstrate that there are objective moral values whose realization can lead to a successful, productive, and happy life.

References

- Achieve (2015) "Employers and College Faculty Report Gaps in Recent Graduates' Preparedness in New National Survey." Available at www.achieve.org/employers-and-college-faculty-report-gaps-recent-graduates%E2%80%99-preparedness-new-national-survey.
- Baumol, W. J., and Blinder, A. S. (2005) *Economics: Principles and Policy*. Mason, Ohio: South-Western.
- Bestor, A. E. (1953) *Educational Wastelands: The Retreat from Learning in Our Public Schools*. Urbana: University of Illinois Press.
- Blanshard, B. (1939) *The Nature of Thought*, Vol. 2. London: George Allen & Unwin.
- Blitzer, R. (2001) *Precalculus*. Upper Saddle River, N.J.: Prentice-Hall.

- Bloom, B. S., ed. (1956) *Taxonomy of Educational Objectives: The Classification of Educational Goals*, Handbook I: Cognitive Domain. New York: David Kay.
- Brinkley, A. (2007) *American History: A Survey*. New York: McGraw Hill.
- Carbaugh, R. J. (2008) *Contemporary Economics*. Armonk, N.Y.: M. E. Sharp.
- Choi, S. P., and Chen, X. (1998) "Toward Better Teaching: Professional Development in 1993–94." Available at nces.ed.gov/pub98/98230.pdf.
- Clayton, G. E. (1995) *Economics: Principles and Practices*. New York: McGraw Hill.
- Cooper, J. M., ed. (1999) *Classroom Teaching Skills*. Boston: Houghton Mifflin.
- Cremin, L. (1961) *The Transformation of the School*. New York: Vintage Books.
- Dewey, J. (1897) "My Pedagogic Creed." *School Journal* 54 (3): 77–80.
- _____ (1915) *The School and Society*. Mineola, N.Y.: Dover Publications.
- _____ (1929) *The Quest for Certainty*. Carbondale, Ill.: Southern Illinois University Press.
- Dressler, I., and Keenan, E. P. (1989) *Integrated Mathematics*, Course I. New York: Amsco School Publications.
- Feltovich, P. J.; Prietula, M. J.; and Ericsson, K. A. (2006) "Studies of Expertise from Psychological Perspectives." Available at www.researchgate.net/publication/200772882_Studies_of_expertise_from_psychological_perspectives.
- Fowers, B. J., and Richardson, F. C. (1996) "Why Is Multiculturalism Good?" *American Psychologist* 51 (6): 609–21.
- Gagne, R. M. (1965) *Conditions of Learning*. New York: Holt, Reinhart and Winston.
- Hirsch, E. D. (2016) *Why Knowledge Matters*. Cambridge, Mass.: Harvard Education Press.
- Hutchins, R. (1953) *Conflict in Education in a Democratic Society*. New York: Harper & Bros.
- Kilpatrick, W. (1925) *Foundations of Method*. New York: Macmillan.
- Klein, M. (1973) *Why Johnny Can't Add*. New York: St. Martin's Press.

- Lewis, L.; Parsad, B.; Carey, N.; Bartfai, N.; Farris, E.; and Smerdon, B. (1999) "Teacher Quality: A Report on the Preparation and Qualifications of Public School Teachers." Available at nces.ed.gov/pubs99/1999080.pdf.
- Lynd, A. (1953) *Quackery in the Public Schools*. Boston: Little Brown.
- Mulroy, D. (2003) *The War Against Grammar*. Portsmouth, N.H.: Boynton/Cook.
- National Assessment of Educational Progress (2017) "The Nation's Report Card." Available at www.nationsreportcard.gov.
- Paterson, I. (1943) *The God of the Machine*. New Brunswick, N.J.: Transaction Publishers.
- Planer, M. G., and Neff, W. L. (1962) *Freedom under Law: A History of the United States*. Milwaukee, Wis.: Bruce Publishing Co.
- Public Agenda (2004) "Teaching Interrupted: Do Discipline Policies in Today's Public Schools Foster the Common Good?" Available at www.publicagenda.org/files/teaching_interrupted.pdf.
- Rand, A. (1966) *Introduction to Objectivist Epistemology*. New York: Meridan.
- Ravitch, D. (1983) *The Troubled Crusade*. New York: Basic Books.
- _____ (1985) "Decline and Fall of Teaching History." *New York Times* (November 17).
- _____ (2003) "A Brief History of Teacher Professionalism." Available at www2.ed.gov/admins/tchrqual/learn/preparingteachers_conference/ravitch.html.
- Rugg, H., and Shumaker, A. (1928) *The Child-Centered School*. New York: World Book.
- Scafidi, B. (2016) "The Dismal Productivity Trend for Public K–12 Schools and How to Improve It." *Cato Journal* 36 (1): 121–41.
- Schweikart, L. (2008) *48 Liberal Lies about American History*. New York: Penguin.
- Staklis, S., and Henke, R. (2013) "Who Considers Teaching and Who Teaches?" Available at nces.ed.gov/pubs2014/2014002.pdf.
- VanDamme, L. (2006) "The Hierarchy of Knowledge: The Most Neglected Issue in Education." *The Objective Standard* 1 (1): 55–81.
- Vedder, R. (2000) *Can Teachers Own Their Own Schools? New Strategies for Educational Excellence*. Oakland, Calif.: The Independent Institute.

- Whitehead, A. N., and Russell, B. (1910) *Principia Mathematica*, Vol. 1. London: Cambridge University Press.
- Wilgoren, J. (2002) "School Cheating Scandal Tests a Town's Values." *New York Times* (February 14).
- YouthTruth (2015) "Most High Schoolers Feel Unprepared for College and Career." Available at www.youthtruthsurvey.org/wp-content/uploads/2016/01/YouthTruth-Learning-From-Student-Voice-College-and-Career-Readiness-2016.pdf.

WHAT LEADS TO SUCCESSFUL SCHOOL CHOICE PROGRAMS? A REVIEW OF THE THEORIES AND EVIDENCE

*Corey A. DeAngelis and
Heidi Holmes Erickson*

There is a large body of research showing many positive benefits of school choice. However, many questions remain on how school choice works. Rigorous school choice experiments can only determine if access to school choice programs alters student outcomes; they cannot confidently identify the specific mechanisms that mediate various outcomes. Two commonly discussed mechanisms in school choice programs, thought to generate positive student outcomes, are (1) an increased access to higher-quality schools and (2) an improved match between schools and students. We examine the existing empirical evidence and arguments for these two primary mechanisms. While there is evidence supporting both mechanisms, no studies are able to isolate the effect of quality schools on student outcomes independent of families selecting schools that match their preferences. Since the majority of this research is descriptive and has limited causal interpretation, theory is essential in guiding interpretation and policy implications. Theory suggests that people make choices based on what they believe to be the best match for their children, and those choices lead to incentives for individual schools

Cato Journal, Vol. 38, No. 1 (Winter 2018). Copyright © Cato Institute. All rights reserved.

Corey A. DeAngelis is a Policy Analyst at the Cato Institute's Center for Educational Freedom, and Heidi Holmes Erickson is a Doctoral Fellow with the Department of Education Reform at the University of Arkansas.

to improve. We conclude with policy recommendations based on our summary of the literature.

School Choice

Families frequently participate in public school choice by selecting where they live. This is known as “Tiebout choice.” Economist Charles Tiebout (1956) posited that families select their homes based on a basket of goods and services, including their schools. Since the decision to move is based on many factors—such as proximity to a city center, pollution levels, commute times, safety, budget constraints, and natural surroundings—moving is a very costly option for a family in order to opt out of their residentially assigned public school. More realistic public school choice comes in the form of publicly funded and privately managed charter schools that usually have specialized missions. State law requires that public charter schools are tuition-free, and, when oversubscribed, most schools use a random lottery to determine which children attend.

Publicly funded private school choice is available to fewer families. The most well-known type of private school choice is school vouchers. Vouchers provide families with public money to attend a private school of their choice. Private school vouchers are often attributed to Milton Friedman (1962), but K–12 education vouchers have existed in the United States since the 19th century in Maine and Vermont. John Stuart Mill (1869) advocated education vouchers just before town-tuition programs were implemented in Maine and Vermont. Other types of private school choice include tax-credit scholarships, tax-credit deductions, and, most recently, Education Savings Accounts (ESAs). While there are slight differences in each of these programs, they all make it less costly for parents to opt out of their residentially assigned public school in order to send their children to the school that better fits their needs.

There is substantial evidence that private choice programs have positive effects for students. A meta-analysis of 19 voucher experiments around the world indicates that private school choice improves student math and reading test scores (Shakeel, Anderson, and Wolf 2016). Out of 20 experimental evaluations of private school choice in the United States, only two find negative impacts on student math and reading test scores (Abdulkadiroglu, Pathak, and Walters 2015; Dynarski et al. 2017). One notable experiment

(Wolf et al. 2013) shows that the D.C. voucher program increased the likelihood of high school graduation by 21 percentage points and one quasi-experiment (Cowen et al. 2013) finds that the Milwaukee Parental Choice Program (MPCP) increased high school graduation rates by 3 percentage points. While research of school choice effects on educational attainment is limited, a review of 12 studies suggests that private and public school choice has a positive effect on student attainment (Foreman 2017).

A review of the experimental and quasi-experimental evidence finds that U.S. private school choice programs reduce criminal activity, increase civic engagement, and increase tolerance of others (DeAngelis 2017b). Another review of the evidence shows that seven out of eight voucher studies conclude that private school choice improves racial integration (Swanson 2017). Furthermore, Egalite's (2013) review finds that 20 out of 21 empirical studies indicate that competitive pressures from school choice programs improve test scores for students who remain in traditional public schools. In addition, more than 20 evaluations (Forster 2016) have found that all of these benefits result in state (e.g., Costrell 2010, Spalding 2014, Trivitt and DeAngelis 2016) and district-level (Scafidi 2012) financial savings.

The question remains: How does school choice lead to these benefits? We examine two possible mechanisms: (1) an increase in the supply of generally better schools, and (2) a better match between educators and students. While these two mechanisms are closely related, and difficult to completely disentangle, we scrutinize the relative strengths of the mechanisms and summarize the current empirical evidence.

We discuss the impact of quality schools on the success of school choice programs, but we are not concerned with defining the absolute measures of school quality. We recognize that measuring school quality is a highly debated issue in education policy, but that is not the aim of this article. Rather, our focus is to discuss the mechanisms for improving student outcomes and to examine the evidence for successful school choice programs.

Mechanisms for Improving Student Outcomes

Within school choice programs, there are a variety of mechanisms that can lead to improved student outcomes. We will focus on two of

the most compelling: (1) market competition to increase the supply of high-quality schools, and (2) improving the match between schools and students.

Market Competition

The first mechanism is largely related to basic economic theory. The traditional public school system in the United States—and around most of the world—consists of a strong public school monopoly financed by taxes (Merrifield 2001). Families pay taxes that finance K–12 public schools whether their children attend public schools or not. If families are dissatisfied with their residentially assigned options, they can opt out of their school only if they move neighborhoods or pay for private school tuition. Meanwhile, they continue paying for public schools indirectly through the tax system. Consequently, there are few incentives for public schools to innovate or respond to families' needs.

For example, imagine if a company could force people to pay for its services regardless of its quality or if individuals choose to purchase its products. As long as the company met the minimum standards set by the government, it would remain profitable without having to respond to the needs of customers. Conversely, when markets function in an open system, competitive pressures lead to quality improvement (Hoxby 2003). School choice programs diminish monopoly power held by traditional public school leaders and, therefore, lead to increased overall quality levels and lower costs (Chubb and Moe 1990, Friedman and Friedman 1990). In other words, market pressures could change the supply of schools by enticing high-quality schools to open and persist while incentivizing low-quality institutions to either improve or close down.

Parental Choice in the Student–School Match

The second mechanism focuses on the ability of families to choose their children's educational institution, which allows for a better match between schools and students. Public K–12 education is a one-size-fits-all system that is unable to serve students' varied needs. Since all children have unique interests, ability levels, desires, and learning styles, an improved student–school match can lead to better student outcomes.

There is a significant theoretical problem with separating the two potential mechanisms and determining which is primarily responsible for generating positive student outcomes. If the student–school match improves, student outcomes will improve. If students have access to better-quality schools, student outcomes will also improve. However, if the quality of schooling options increases, the likelihood of a strong student–school match also increases.

To better understand the causal stories of the two primary mechanisms, it is helpful to think of different definitions of quality. If the definition of quality is unique to each individual, we could say that the school selection itself—the student–school match—is the definition of quality. Indeed, since the match is likely made based on attributes that are unique to each family situation, school quality can increase without a change in observable metrics such as student test scores. For example, many children may largely benefit from vocational training or French immersion, which may not necessarily translate to higher standardized test scores. In addition, the perceived match—based on income, race, gender, test scores, ability levels, family culture, distance to home, school safety, learning styles, and other individual interests—leads to competitive pressures that alter the quality levels of current and future schools existing in the market.

In theory, school quality could remain stagnant while student test scores improve. However, if overall school quality is defined as test score gains, and student outcomes are measured by the test scores, there is an obvious connection between the two mechanisms. Student matching could lead to improved test scores and, by definition, higher-quality schools. Alternatively, higher-quality schools opening up could lead to more and better opportunities for matching between schools and students. Nonetheless, if quality is objectively measured, we find ourselves with a classic chicken-and-the-egg predicament. Did the match boost measures of quality in the schools, or did the robust supply of quality schools allow for better matching? Concurrently, did a robust supply of quality schools improve student outcomes, regardless of the quality of the match?

Existing Literature

While there is an abundance of literature on the effects of school choice programs on student outcomes—such as test scores,

graduation rates, civic skills, and parental satisfaction—it is particularly difficult to confidently determine the precise program mechanisms at play. Even the best school choice experiments do not provide insight because experimental evaluations treat programs as black boxes—that is, they simply determine the effects of programs without explaining why. Randomly assigning children to public or private schools can tell us about the *average* differences in student outcomes but cannot tell us exactly why one group outperforms the other. We must rely on theory and a few descriptive analyses that have attempted to peer into the black box of school choice experiments in the United States.

Competitive Effects of School Choice on Quality

The most robust literature examines the competitive effects of school choice programs. These studies look at the effect competitive pressures created by school choice programs have on students who remain in traditional public schools. Since those students come from families that do not actively select schools, we can be more confident that the effects are not exclusively from improved student–school matches. Traditional public schools face a financial incentive to keep children in their schools, and they must improve quality by changing instructional practices (Rouse et al. 2013). These studies indicate that increasing the supply of schools in the market has a positive effect on students.

Egalite’s (2013) review of the competitive effects literature shows that 20 of 21 studies find positive effects of private school choice programs on student test scores in traditional public schools. Three other studies have been released since Egalite’s (2013) review, and 23 of 24 evaluations have found positive competitive effects for student test scores in district schools (e.g., Chakrabarti 2013, Egalite 2014, Egalite 2016). None of the studies found negative effects. A recent study on this topic finds that nearby public charter schools in New York City lead to increases in student math and reading achievement in local district schools (Cordes 2017). However, one must not overlook the fact that, although the affected district school students did not have freedom of choice, the competitive effects are driven by self-interested schooling selections made by families using school choice programs. The selections made by those families must be made by their subjective definitions of quality. Those decisions

could be made based on academic measures, such as standardized test scores, or more subjective measures, such as school mission, culture, and even safety (Stewart and Wolf 2016).

Wolf and Hoople (2009) descriptively look at school factors that explain voucher gains in Washington, D.C. They find that more advantaged peers, responsible teachers, and more time-consuming homework may increase academic achievement of voucher recipients in the nation's capital. All three of these attributes are commonly associated with what one considers to be greater school quality. In a similar study, Berends et al. (2010) find no clear charter school effect in three states, but find that high-quality instructional conditions, such as teacher quality and a focus on academics, explain gains in math test scores. Another study further supports the quality mechanism, finding that high-performing charter schools have longer school days, comprehensive behavioral policies, intense tutoring, teacher feedback and coaching, and data-driven instructional practices (Gleason 2016). Nonetheless, this same empirical analysis finds that successful charter schools are more likely to have a mission that prioritizes student academic achievement (Gleason 2016). While this attribute is commonly associated with higher school quality, one could also make the case that a mission based on academic achievement could improve the match between schools and students. If the school has a clear mission based on academics, families interested in shaping math and reading test scores could be more likely to choose that institution. In addition, Hoxby (2000) finds that school choice enhances competition between schools and increases the demand for high-quality teachers.

Student–School Match Mechanism

It is difficult to rigorously assess the quality of a student–school match because it requires knowing the preferences and needs of students, the quality of the school they previously attended, and the quality of the school of choice. Focus groups and surveys are important in understanding what parents want when shopping for a school and if their selected school meets those standards. If parents have clear preferences, select schools based on those preferences, and experience better outcomes, it follows that school choice succeeds by allowing for a better match between schools and students.

Parents consider a variety of factors that are specific to their circumstances when participating in school choice, and they make tradeoffs among their preferences based on their needs. Parents participating in choice programs report considering a variety of factors, such as curriculum (Stewart et al. 2009, Stewart et al. 2010), better academics than previous schools (Kelly and Scafidi 2013, Catt and Rhinesmith 2016), test scores (Lincove, Cowen, and Imbrogno 2016), class size (Stewart et al. 2009, Catt and Rhinesmith 2016), individual attention/better learning environment (Kelly and Scafidi 2013, Catt and Rhinesmith 2016), school safety (Stewart et al. 2009, Kelly and Scafidi 2013), religious or moral instruction (Stewart et al. 2009, Kelly and Scafidi 2013, Catt and Rhinesmith 2016), sport programs (Stewart et al. 2010; Lincove, Cowen, and Imbrogno 2016), school convenience (Stewart et al. 2010), school reputation (Stewart et al. 2010), and the child's preference (Stewart et al. 2010).

It is clear that parents have preferences and select schools based on them. Stewart et al. (2009) and Stewart et al. (2010) conduct focus groups for parents participating in school voucher programs in D.C. and Milwaukee and find that parents seek educational institutions that fit their children's needs better than their traditional public schools. For example, Stewart et al. (2009) find that many parents in the nation's capital choose private schools for increased safety and report that they worry less about their children's safety at their schools. There is also evidence that students experience the desired outcomes that parents want from a better school match. Hastings and Weinstein (2008) take advantage of a natural experiment and show that lower-income parents receiving direct information on academic performance are more likely to exercise public school choice as a means toward academic achievement. Consequently, children of informed low-income parents experience increased test scores. In addition, a recent study by DeAngelis (2017a) compares open-enrollment (choice) charters to district-conversion (nonchoice) charters in Arkansas and suggests that parental satisfaction is higher in choice schools—even after controlling for family and student background characteristics.

Some of the previously mentioned studies indicate that shaping the supply of high-quality schools improves student outcomes. It is difficult to measure the quality of a student-school match because it depends on unobservable student and school characteristics. For example, one of the studies finds that choice students who are around

advantaged peers (based on income and prior achievement levels) experience increased academic achievement (Wolf and Hoople 2009). However, these same students could have matched with students that share the same interests. Perhaps the students are diverse based on socioeconomic status, but similar on learning styles and academic interests. While studies such as this one show that diversity of household income levels may be beneficial, the observable characteristic may be correlated with important unobservable traits such as family culture, student curiosity, learning style, and long-term goals.

Theoretically, it may be that student–school matching and high-quality schools are necessary, but not sufficient, to elicit the positive outcomes seen in school choice programs; both mechanisms are likely essential in school choice. Notably, the observational study designs that even the best social scientists are limited by may never allow us to separate the two mechanisms with quantitative analyses. If, for example, parents matched their children to schools based on a mix of academic rigor, school culture, safety, and moral education, how would one begin to assess the match? The task would be nearly impossible for researchers to perform for each individual family and child, especially since experiments require grouping people to make causal claims (Federer 1955; Rossi, Lipsey, and Freeman 2003). These studies can only descriptively tell us how families choose schools, their preferences, and the types of schools they select.

Policy Implications

The empirical findings on school choice mechanisms are mixed overall; in part because the two primary mechanisms are acutely connected and difficult to disentangle. Based on the empirical evidence and the interconnectedness of the two mechanisms, we cannot determine with certainty which one is principally responsible for the positive outcomes of schools choice. Because of the severe limitations of the existing empirical analyses, we should be cautious in using them to design school choice programs.

Policymakers trying to design an effective school choice program could look at limited evidence on school quality as the mechanism and hastily conclude that regulated school choice is the best path forward since successful choice schools often have characteristics that are associated with high school quality such as increased seat time, time-consuming homework, and qualified teachers. If government officials

could feasibly limit the school choices of families only to high-quality institutions, should they not do so?

The answer to this question is particularly unclear for four fundamental reasons: (1) observational analyses, by definition, can only be based on observable characteristics, while parents match their children to schools based on numerous observable and unobservable characteristics; (2) even if we could determine what the “secret sauce” is made out of today, the factors that lead to educational success likely differ across locations and students and change over time; (3) alternative evidence suggests that attempting to control the quality of the supply of schools reduces overall school quality; and (4) there is a growing body of school choice evidence indicating a disconnect between short-term observable measures and arguably more important long-term student outcomes.

Ironically, in failing to trust families with the decision of selecting a school that meets their children’s needs, policymakers inadvertently lower the amount of available quality schools. Sude, DeAngelis, and Wolf (2017) find that higher-quality private schools are less likely to participate in highly regulated voucher programs. Further, only a third of the private schools in Louisiana participate in the most regulated program in the study—the Louisiana Scholarship Program. Regulations may very well lead to fewer choices overall since they serve as a significant cost for participating schools. Similarly, Kisida, Wolf, and Rhinesmith (2015) find that the biggest concern for leaders of schools participating in the Louisiana Scholarship Program is the likelihood of future regulations. Further, DeAngelis and Burke (2017) find that private schools in more highly regulated voucher programs are less likely to be specialized. Evaluations of the Louisiana voucher program were also the first experiments to find statistically significant and large negative effects on student achievement (Abdulkadiroglu, Pathak, and Walters 2015; Mills and Wolf 2017).

The ability of families to match their children to an appropriate schools is obviously at least partially related to whether they have the information necessary to make decisions that would maximize each child’s utility. While parents may not currently be education experts, recent evidence from online search behavior finds that school choice programs increase the amount of information gathered on differences in quality across schools (Lovenheim and Walsh 2017). Other research has also found that low-income parents gather

information about schools when given the opportunity to choose (Kelly and Scafidi 2013; Teske, Fitzpatrick, and Kaplan 2007). However, even if the information held by parents is less than perfect, the information held by bureaucrats sitting in offices, hundreds of miles away, is also imperfect. Government officials cannot possibly know the situation of each individual family (Hayek 1945), and, even if they did, they are forced to decide what schools are “high-quality” using uniform measures (Buchanan and Tollison 1984). Since all children are inherently unique, uniform policies are likely to fail them.

As Greene (2017) points out, there is a growing literature indicating that short-term changes in student test scores do not necessarily translate to long-term outcomes. For example, some studies on Boston charter schools (Angrist et al. 2016), Harlem Promise Academy (Dobbie and Fryer 2014), and SEED boarding charter in the District of Columbia (Unterman et al. 2016) find substantial test score gains with no increase in high school graduation rates. On the other hand, the voucher programs in Washington, D.C. (Wolf et al. 2013) and Milwaukee (Cowen et al. 2013) produced little or no test score gains with large increases in graduation rates. In other words, improved matches between schools and students can be a good thing even if we do not observe any effects on test scores or related metrics. Consequently, regulating school choices based on the state’s preferred accountability measure—standardized test scores—may very well harm other student outcomes that individuals and society care about.

Despite the empirical difficulties of disentangling school choice mechanisms, we argue that both underlying mechanisms lead to similar policy implications. If the primary mechanism is matching, decisionmakers ought to give families as many choices as possible so that parents could match their children to schools that best fit their needs, whatever those needs may be. An ESA available to all families, regardless of income or ability levels, would allow parents to customize the educational environment for their children to the best of their abilities. Importantly, ESAs allow parents to fit student needs for schooling, tutoring, textbooks, and even college. A universal program would generate enough demand for robust market entry in the long run, meaning more choices for all families. If parents do not perceive that certain schools or services will be

appropriate for their children, they will not choose them—enticing schools to improve or force them to close down. The schools that are a quality match for many children will be financially rewarded and expand in the long run.

If the primary school choice mechanism is the supply of high-quality schools, we should allow the market to determine which institutions are high quality. The choices of individual parents, rather than bureaucrats, can determine which schools remain open and which ones close. When public officials choose a uniform measure, such as school test scores, they must determine which level is appropriate for which students. Since student ability levels are diverse, the uniform measure would fail, and since test scores are, at best, a crude proxy for lifelong success, focusing on test scores may result in harming students who would have otherwise benefitted from marginally more diverse education. A universally accessible ESA would allow for robust market entry and customization that would allow individual families to choose high-quality educations. Stronger influxes of demand through a universal school choice program and price differentiation, generated through the ESA, would allow greater competition in the education market to improve student outcomes. Price differentiation sends signals to high- and low-quality schools alike, giving them the information and incentives necessary to provide the best possible K–12 educational experience to all children.

Conclusion

It is likely that we may not be able to empirically disentangle the mechanisms of school choice. Descriptive empirical analyses, by definition, rely on the use of observable characteristics. If we accept the assumption that parents choose schools that are fitted for their children, and that parents want to improve their children's outcomes, we must also accept that the resulting match leads to higher overall school quality levels.

Since the answer to this question cannot be resolved descriptively, we must use sound theory. People make choices based on what they perceive as the best match for their children, and those choices create incentives for individual schools to improve. The supply of quality schools then improves because families choose educational products that best fit their needs. Regardless of which mechanism is

the most important, ESA programs that are accessible to all children, regardless of incomes, abilities, or other background characteristics, offer the best opportunity for a robust market to thrive where the supply of quality schools can increase and families are able find the best educational match.

References

- Abdulkadiroglu, A.; Pathak, P. A.; and Walters, C. R. (2015) “Free to Choose: Can School Choice Reduce Student Achievement?” NBER Working Paper No. 21839. National Bureau of Economic Research, Cambridge, Mass.
- Angrist, J. D.; Cohodes, S. R.; Dynarski, S. M.; Pathak, P. A.; and Walters, C. R. (2016) “Stand and Deliver: Effects of Boston’s Charter High Schools on College Preparation, Entry, and Choice.” *Journal of Labor Economics* 34 (2): 275–318.
- Berends, M.; Goldring, E.; Stein, M.; and Cravens, X. (2010) “Instructional Conditions in Charter Schools and Students’ Mathematics Achievement Gains.” *American Journal of Education* 116 (3): 303–35.
- Buchanan, J. M., and Tollison, R. D., eds. (1984) *The Theory of Public Choice—II*. Ann Arbor: University of Michigan Press.
- Chakrabarti, R. (2013) “Vouchers, Public School Response, and the Role of Incentives: Evidence from Florida.” *Economic Inquiry* 51 (1): 500–26.
- Chubb, E. J., and Moe, T. M. (1990) *Politics, Markets, and America’s Schools*. Washington: The Brookings Institution.
- Cordes, S. A. (2017) “In Pursuit of the Common Good: The Spillover Effects of Charter Schools on Public School Students in New York City.” *Education Finance and Policy*. Available at https://doi.org/10.1162/edfp_a_00240.
- Costrell, R. M. (2010) “The Fiscal Impact of the Milwaukee Parental Choice Program: 2010–2011 Update and Policy Options.” SCDP Milwaukee Evaluation Report No. 22. School Choice Demonstration Project.
- Cowen, J. M.; Fleming, D. J.; Witte, J. F.; Wolf, P. J.; and Kisida, B. (2013) “School Vouchers and Student Attainment: Evidence from a State Mandated Study of Milwaukee’s Parental Choice Program.” *Policy Studies Journal* 41 (1): 147–68.

- DeAngelis, C. A. (2017a) “Does Choice Matter for School Choice? An Instrumental Variables Analysis of the Effect of Choice on Parental Satisfaction in Charter Schools.” *Journal of School Choice*. Available at <https://doi.org/10.1080/15582159.2017.1408003>.
- _____ (2017b) “Do Self-Interested Schooling Selections Improve Society? A Review of the Evidence.” *Journal of School Choice* 11 (4): 546–58.
- DeAngelis, C. A., and Burke, L. (2017) “Does Regulation Induce Homogenization? An Analysis of Three Voucher Programs in the United States.” EDRE Working Paper No. 2017-14. University of Arkansas, Department of Education Reform.
- Dobbie, W., and Fryer, R. G. Jr. (2014) “The Impact of Attending a School with High-Achieving Peers: Evidence from the New York City Exam Schools.” *American Economic Journal: Applied Economics* 6 (3): 58–75.
- Dynarski, M.; Rui, N.; Webber, A.; and Gutmann, B. (2017) “Evaluation of the DC Opportunity Scholarship Program: Impacts after One Year.” NCEE 2017-4022. National Center for Education Evaluation and Regional Assistance.
- Egalite, A. J. (2013) “Measuring Competitive Effects from School Voucher Programs: A Systematic Review.” *Journal of School Choice* 7 (4): 443–64.
- _____ (2014) “Competitive Impacts of Means-Tested Vouchers on Public School Performance: Evidence from Louisiana and Indiana.” PEPG Working Paper 14. Program on Education Policy and Governance, Harvard Kennedy School.
- _____ (2016) “The Competitive Effects of the Louisiana Scholarship Program on Public School Performance.” Louisiana Scholarship Program Evaluation Report No. 4. School Choice Demonstration Project, University of Arkansas.
- Federer, W. T. (1955) *Experimental Design*. New York: Macmillan.
- Foreman, L. M. (2017) “Educational Attainment Effects of Public and Private School Choice.” EDRE Working Paper 2017-16. University of Arkansas, Department of Education Reform.
- Forster, G. (2016) “A Win-Win Solution: The Empirical Evidence on School Choice.” EdChoice. Available at www.edchoice.org/wp-content/uploads/2016/05/2016-5-Win-Win-Solution-WEB.pdf.
- Friedman, M. (1962) *Capitalism and Freedom*. Chicago: University of Chicago Press.

- Friedman, M., and Friedman, R. (1990) *Free to Choose: A Personal Statement*. Orlando, Fla.: Houghton Mifflin Harcourt.
- Gleason, P. M. (2016) “What’s the Secret Ingredient? Searching for Policies and Practices That Make Charter Schools Successful.” Mathematica Policy Research. Available at www.mathematica-mpr.com/our-publications-and-findings/publications/whats-the-secret-ingredient-searching-for-policies-and-practices-that-make-charter-schools.
- Greene, J. P. (2017) “Evidence for the Disconnect between Changing Test Scores and Changing Later Life Outcomes.” *Education Next*. Available at <http://educationnext.org/evidence-for-the-disconnect-between-changing-test-scores-and-changing-later-life-outcomes>.
- Hastings, J. S., and Weinstein, J. M. (2008) “Information, School Choice, and Academic Achievement: Evidence from Two Experiments.” *Quarterly Journal of Economics* 123 (4): 1373–414.
- Hayek, F. A. (1945) “The Use of Knowledge in Society.” *American Economic Review*: 35 (4): 519–30.
- Hoxby, C. (2000) “Would School Choice Change the Teaching Profession?” NBER Working Paper No. 7866. National Bureau of Economic Research, Cambridge, Mass.
- _____ (2003) “School Choice and School Productivity. Could School Choice Be a Tide That Lifts All Boats?” In C. Hoxby (ed.) *The Economics of School Choice*, 287–342. Chicago: University of Chicago Press.
- Kelly, J. P., and Scafidi, B. (2013) “More Than Scores: An Analysis of Why and How Parents Choose Private Schools.” The Friedman Foundation for Educational Choice. Available at www.edchoice.org/wp-content/uploads/2015/07/More-Than-Scores.pdf.
- Kisida, B.; Wolf, P. J.; and Rhinesmith, E. (2015) “Views from Private Schools: Attitudes about School Choice Programs in Three States.” American Enterprise Institute for Public Policy Research. Available at www.aei.org/publication/views-private-schools-attitudes-school-choice-programs-three-states.
- Lincove, J.; Cowen, J. M.; and Imbrogno, J. (2016) “What’s in Your Portfolio? How Parents Rank Traditional Public, Private, and Charter Schools in Post-Katrina New Orleans’ Citywide System of School Choice.” *Education Finance and Policy*, https://doi.org/10.1162/EDFP_a_00222.

- Lovenheim, M. F., and Walsh, P. (2017) “Does Choice Increase Information? Evidence from Online School Search Behavior.” NBER Working Paper No. 23445. National Bureau of Economic Research, Cambridge, Mass.
- Merrifield, J. D. (2001) *The School Choice Wars*. Lanham, Md: R&L Education.
- Mill, J. S. (1869) *On Liberty*. London: Longmans, Green, Reader, and Dyer.
- Mills, J. N., and Wolf, P. J. (2017) “Vouchers in the Bayou: The Effects of the Louisiana Scholarship Program on Student Achievement after 2 Years.” *Educational Evaluation and Policy Analysis* 39 (3): 464–84.
- Rossi, P. H.; Lipsey, M. W.; and Freeman, H. E. (2003) *Evaluation: A Systematic Approach*. Thousand Oaks, Calif.: SAGE Publications.
- Rouse, C. E.; Hannaway, J.; Goldhaber, D.; and Figlio, D. (2013) “Feeling the Florida Heat? How Low-Performing Schools Respond to Voucher and Accountability Pressure.” *American Economic Journal: Economic Policy* 5 (2): 251–81.
- Scafidi, B. (2012) “The Fiscal Effects of School Choice Programs on Public School Districts.” *EdChoice*. Available at <http://files.eric.ed.gov/fulltext/ED529881.pdf>.
- Shakeel, M.; Anderson, K. P.; and Wolf, P. J. (2016) “The Participant Effects of Private School Vouchers across the Globe: A Meta-Analytic and Systematic Review.” EDRE Working Paper No. 2016-07. University of Arkansas, Department of Education Reform.
- Spalding, J. (2014) “The School Voucher Audit: Do Publicly Funded Private School Choice Programs Save Money?” *EdChoice*. Available at www.edchoice.org/wp-content/uploads/2015/07/The-School-Voucher-Audit-Do-Publicly-Funded-Private-School-Choice-Programs-Save-Money.pdf.
- Stewart, T., and Wolf, P. (2016) *The School Choice Journey: School Vouchers and the Empowerment of Urban Families*. New York: Palgrave MacMillan.
- Stewart, T.; Wolf, P.; Cornman, S. Q.; McKenzie-Thompson, K.; and Butcher, J. (2009) “Family Reflections on the District of Columbia Opportunity Scholarship Program. Final Summary Report.” School Choice Demonstration Project, University of Arkansas.
- Stewart, T.; Lucas-McLean, J.; Jensen, L. I; Fetzko, C.; Ho, B.; and Segovia, S. (2010) “Family Voices on Parental School Choice in

- Milwaukee: What Can We Learn from Low-Income Families?" SCDP Milwaukee Evaluation Report No. 19. School Choice Demonstration Project, University of Arkansas.
- Sude, Y.; DeAngelis, C. A.; and Wolf, P. J. (2017) "Supplying Choice: An Analysis of School Participation Decisions in Voucher Programs in Washington, DC, Indiana, and Louisiana." *Journal of School Choice* 11 (4): 1–26.
- Swanson, E. (2017) "Can We Have It All? A Review of the Impacts of School Choice on Racial Integration." *Journal of School Choice* 11 (4): 507–26.
- Teske, P.; Fitzpatrick, J.; and Kaplan, G. (2007) "Opening Doors: How Low-Income Parents Search for the Right School." Center on Reinventing Public Education. Daniel J. Evans School of Public Affairs, University of Washington.
- Tiebout, C. M. (1956) "A Pure Theory of Local Expenditures." *Journal of Political Economy* 64 (5): 416–24.
- Trivitt, J. R., and DeAngelis, C. A. (2016) "The Fiscal Effect of Eliminating the Louisiana Scholarship Program on State Education Expenditures." EDRE Working Paper No. 2016-06. University of Arkansas, Department of Education Reform.
- Unterman, R.; Bloom, D.; Byndloss, D.; and Terwelp, E. (2016) "Going Away to School: An Evaluation of SEED DC." MDRC. Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2821816.
- Wolf, P. J., and Hoople, D. S. (2006) "Looking Inside the Black Box: What School Factors Explain Voucher Gains in Washington, DC?" *Peabody Journal of Education* 81 (1): 7–26.
- Wolf, P. J.; Kisida, B.; Gutmann, B.; Puma, M.; Eissa, N.; and Rizzo, L. (2013) "School Vouchers and Student Outcomes: Experimental Evidence from Washington, DC." *Journal of Policy Analysis and Management* 32 (2): 246–70.

EXTENDING THE ECONOMIC FREEDOM OF THE WORLD INDEX TO THE COLD WAR ERA

Ryan H. Murphy and Robert A. Lawson

The Economic Freedom of the World (EFW) index was first published by the Fraser Institute in 1996 and featured data in five-year increments from 1975 to 1995 (Gwartney, Block, and Lawson 1996). Subsequent editions of the EFW index (e.g., Gwartney, Lawson, and Hall 2016) have included data for 1970 and annually since 2000. The EFW index has become a useful tool for empirical researchers studying how institutions consistent with economic freedom correlate with various socioeconomic indicators (Hall and Lawson 2014); this is especially true with respect to economic growth (De Haan, Lundstrom, and Sturm 2006).

This article uses newly gathered and available data and autoregressive methods to create an economic freedom index for the 1950s and 1960s for up to 95 countries. The resulting index allows not only for a longer time series but also for a larger sample of countries than has been previously available.

It is critically important that the new index we develop be reasonably consistent with the existing EFW index. For this reason, we use two methods to mesh the new data we have collected with the existing EFW data. First, we collect data for 1950–80, in five-year intervals, using newly gathered and available data to create a new index. The overlap from 1970 to 1980 will later allow us to parameterize weights.

Cato Journal, Vol. 38, No. 1 (Winter 2018). Copyright © Cato Institute. All rights reserved.

Ryan H. Murphy is a Research Assistant Professor, and Robert A. Lawson holds the Jerome M. Fullinwider Endowed Centennial Chair in Economic Freedom, at the O'Neil Center for Global Markets and Freedom, Southern Methodist University. Lawson is also an author of the *Economic Freedom of the World* report.

This new index (EFWNEW) is designed to mirror the structure of the current EFW index but because of limited data includes only 8 components rather than 24.¹ Second, we use the 1990–2013 EFW ratings to create an autoregressive model to back cast the EFW data to 1950. Because there are more countries available after 1990, this allows us to create index estimates (EFWHAT) for a much larger set of nations than is available in the current EFW index. Finally, the EFWNEW and EFWHAT ratings for 1970, 1975, and 1980 are regressed against the actual EFW ratings. The coefficients from this regression are then used to generate economic freedom ratings for 1950–65 based on the EFWNEW and the EFWHAT ratings for 1950–65.

The development and importance of institutions, variously defined, are an important topic in economic history (see, e.g., Chor 2005; Johnson and Koyama 2017). But most literature evaluating the historical evolution of liberalism is narrative, not quantitative. A notable exception is De La Escosura (2016), who mirrors the EFW index back to 1950 for OECD countries. A working paper by Murphy and Stansel (2017) attempts to do so for individual U.S. states by decade in the 19th century. The broader question of how institutions play a role in the very long run has been addressed, in part, using measures such as colonial deaths (Acemoglu, Johnson, and Robinson 2001) and legal origins (La Porta et al. 1997; Glaeser and Shleifer 2002; La Porta, Lopez-De-Silanes, and Shleifer 2008). Narratives with a focus on the institutional history and economic freedom in the United States include Higgs (1987) and Holcombe (2002), and historical property rights in the United States for women can be found in Geddes and Tennyson (2013) and Lemke (2016). Narratives regarding economic freedom and the development of liberal market institutions in the international context can be found for Australia (Berg 2015), China (Wang and Coase 2012), Czech Republic (Sima and Nikodym 2015), Denmark (Kurrild-Klitgaard 2015), Georgia (Burakova and Lawson 2014), Guatemala (Marroquin and Thomas 2015), India (Manish et al. 2015), Korea (Choi and Yoon 2016), Lebanon (Mardini 2015), Mexico (Kuchar 2016), Poland (Mateusz 2015), Spain (Fradejas 2015), the Spanish colonies (Hough and Grier 2015), Venezuela (Faria and Filardo 2015), and former Yugoslav nations (Prokopijević and Tasić 2015). A general narrative regarding

¹For a complete listing of the areas, components, and subcomponents of the 2016 EFW index, see Gwartney, Lawson, and Hall (2016: 3).

economic freedom around the world at approximately the same years we cover here is found in Yergin and Stanislaw (1998). Our data we hope can complement these narratives and provide more opportunities for economists to test hypotheses in a historical context.

Constructing a New Economic Freedom Index (EFWNEW), 1950–80

We identified eight variables that mirror the components of the current EFW index for a relatively large number of countries for 1950–80. Table 1 lists these eight components.

Table 2 shows the new index (EFWNEW) ratings for all countries from 1950–80 based only on these eight components. The entire dataset, including all the components, is available from the authors. A description of each variable and the method for computing the 0–10 index rating is described below.

Component 1A: Government Consumption. This component is government consumption as a share of total consumption from the national accounts. The rating is equal to $(V_{\max} - V_i) / (V_{\max} - V_{\min})$ multiplied by 10. V_i is the country's actual government consumption as a proportion of total consumption, while V_{\max} and V_{\min} were set at 40 and 6 respectively. Data were gathered from the *World Development Indicators*, *Penn World Tables*, and Summers and

TABLE 1
AREAS AND COMPONENTS OF A NEW ECONOMIC
FREEDOM INDEX

- Area 1. Size of Government
 - A. Government Consumption
 - Area 2. Legal System and Property Rights
 - A. Judicial Independence
 - Area 3. Sound Money
 - A. Average Inflation (previous 5 years)
 - B. Standard Deviation of Inflation (previous 5 years)
 - C. Foreign Currency Restrictions
 - Area 4. Freedom to Trade Internationally
 - A. Trade Openness
 - B. Import Duties
 - C. Capital Controls
-

TABLE 2
EIGHT-VARIABLE ECONOMIC
FREEDOM INDEX (EFWNEW), 1950-65

Country	1950	1955	1960	1965
Algeria				1.72
Argentina	4.09	4.26	5.09	4.77
Australia	6.91	6.35	7.21	8.41
Austria	5.90	5.72	7.79	7.83
Belgium	6.27	6.70	8.37	8.19
Benin			2.01	1.91
Bolivia	3.90	3.89	5.63	6.50
Brazil	4.71	4.37	5.31	4.83
Burkina Faso			4.82	4.72
Burndi				3.47
Cameroon			2.80	5.87
Canada	7.43	8.64	8.89	8.74
Central African Republic			1.58	1.25
Chad			1.69	1.64
Chile	4.00	4.17	5.56	5.59
China		2.82	2.31	2.69
Colombia	4.97	4.20	6.70	6.58
Congo, Dem. Rep.				3.49
Congo, Rep.			2.80	2.08
Costa Rica	6.06	7.97	7.94	7.08
Côte d'Ivoire			2.77	3.90
Cyprus		7.40	6.51	5.56
Denmark	6.05	5.76	7.55	7.07
Dominican Republic	3.69	5.08	4.86	4.16
Ecuador	5.67	5.98	6.18	6.15
Egypt	5.25	3.52	3.20	2.89
El Salvador	6.78	6.41	6.68	4.76
Ethiopia	2.95	3.72	3.44	2.56
Finland	6.05	5.82	7.03	7.65
France	6.22	5.36	5.28	5.40
Gabon			2.59	4.19
Gambia, The				4.46
Germany	5.91	5.34	7.76	7.85
Ghana		3.31	4.02	2.59
Greece	5.40	5.59	6.30	5.89
Guatemala	6.36	6.32	6.72	4.99

(Continued)

TABLE 2 (Continued)
EIGHT-VARIABLE ECONOMIC
FREEDOM INDEX (EFWNEW), 1950-65

Country	1950	1955	1960	1965
Guinea			3.03	2.50
Guyana		2.75	3.23	3.38
Haiti		4.45	4.71	4.73
Honduras	7.61	6.91	6.86	6.24
Iceland		7.18	7.84	7.65
India	6.25	5.03	5.87	5.28
Indonesia			3.79	4.29
Iran	4.67	4.31	4.12	4.53
Ireland	6.35	5.68	5.81	6.01
Israel	4.80	4.20	5.32	5.25
Italy	7.25	5.68	6.66	6.98
Jamaica		6.78	6.61	9.03
Japan		5.76	6.21	6.74
Jordan		1.39	1.34	3.14
Kenya			6.96	5.75
Korea, South		2.45	3.16	4.05
Libya				2.66
Luxembourg	6.62	6.95	9.33	9.24
Madagascar			2.87	4.22
Malawi				2.56
Malaysia	7.17	5.73	6.49	5.90
Mali			2.68	2.04
Malta			5.35	5.50
Mauritania			1.30	1.58
Mauritius			5.97	7.24
Mexico	5.94	6.33	6.38	6.88
Morocco	5.54	5.22	6.88	3.98
Myanmar	5.57	4.71	3.97	2.73
Nepal			2.77	4.28
Netherlands	5.25	5.66	6.81	7.01
New Zealand	6.62	5.61	5.76	5.74
Nicaragua	5.24	4.64	5.10	3.66
Niger			2.78	4.54
Nigeria		4.62	5.81	4.27
Norway	7.23	5.68	6.74	6.58
Pakistan	3.81	3.54	3.44	4.07

(Continued)

TABLE 2 (Continued)
EIGHT-VARIABLE ECONOMIC
FREEDOM INDEX (EFWNEW), 1950-65

Country	1950	1955	1960	1965
Panama	6.83	7.17	7.27	7.28
Paraguay	1.00	3.66	5.08	5.26
Peru	6.00	6.90	7.34	6.96
Philippines	6.20	5.11	4.84	6.47
Portugal	5.34	6.03	6.42	6.23
Romania			2.96	1.99
Rwanda				2.41
Senegal			2.60	2.82
Sierra Leone			4.63	4.87
Singapore			8.38	7.93
South Africa	5.47	6.31	6.33	6.14
Spain	4.41	4.11	5.23	5.51
Sri Lanka	7.99	5.71	4.36	4.20
Sweden	7.18	6.04	6.98	6.86
Switzerland	8.66	8.78	8.90	8.75
Syria	3.80	5.56	6.51	4.88
Taiwan	0.52	1.70	3.44	5.26
Tanzania				1.44
Thailand	5.45	5.59	5.17	5.15
Togo			3.48	3.49
Trinidad & Tobago		6.34	6.44	6.46
Tunisia			2.76	1.86
Turkey	5.12	4.71	4.72	5.21
Uganda				4.80
United Kingdom	5.57	6.40	6.82	6.80
United States	8.94	7.41	8.73	8.75
Uruguay	5.06	6.33	5.85	5.22
Venezuela	5.85	5.54	5.83	6.26
Zambia				3.47
Zimbabwe				3.50

Heston (1984). This component corresponds to Component 1A in the EFW index.

Component 2A: Judicial Independence. Data from Linzer and Staton (2015) on judicial independence were used to construct 0 to 10 ratings, by simply multiplying their 0 to 1 ratings by 10. This component corresponds to Component 2A in the EFW index.

Component 3A: Average Inflation; and 3B: Standard Deviation of Inflation. Both 3A and 3B use data for the previous five years (e.g., 1955 corresponds to 1951–55, except for 1950, which is 1948–50). The rating is equal to $(V_{\max} - V_i) / (V_{\max} - V_{\min})$ multiplied by 10. V_i represents the actual value (the average inflation rate for 3A or the standard deviation of inflation for 3B). The values for V_{\min} and V_{\max} were set at 0 percent and 50 percent for average inflation, and 0 percent and 25 percent for the standard deviation of inflation. The data source is “Consumer Price Index, All Items, Percent Change, Corresponding Period Previous Year, Percent” in *International Financial Statistics*. These components correspond to Components 3C and 3B in the EFW index, respectively.

Component 3C: Foreign Currency Restrictions. This component measures two dimensions related to foreign currency restrictions: “Free Ownership of Currency within Country” and “Bank Balances Abroad.” The rating was assigned 0 if “no” for both, 5 if “yes” for one, and 10 if “yes” for both. Data are from various issues of *Pick’s Currency Yearbook*. This component corresponds to Component 3D in the EFW index.

Component 4A: Trade Openness. This component is a binary variable (0/10) of trade openness as compiled by Sachs and Warner (1995). This component has no exact corresponding component in the EFW index, but broadly corresponds with EFW Area 4.

Component 4B: Import Duties. This component is “Import Duties Over Imports” as compiled by Clemens and Williamson (2004). The rating is equal to $(V_{\max} - V_i) / (V_{\max} - V_{\min})$ multiplied by 10. V_i is the country’s actual government consumption as a proportion of total consumption, while V_{\max} and V_{\min} were set at 30 percent and 0 percent respectively. This component corresponds to Component 4A in the EFW index.

Component 4C: Capital Controls. This component is based on two dimensions associated with capital controls: “Free Ownership of Foreign Securities” and “Free Export of National Currency.” The rating was assigned 0 if “no” for both, 5 if “yes” for one, and 10 if

“yes” for both. Data are from various issues of *Pick’s Currency Yearbook*. This component corresponds to Components 4Di and 4Dii in the EFW index.

Back Casting the EFW Index to 1950 to Create EFWHAT

The second approach we use is to estimate the following autoregressive models for each of the five EFW areas:

$$(1) \text{ AREA}_t = \beta_0 + \beta_1 \text{ AREA}_{t+5} + \beta_2 \text{ AREA}_{t+10} + \varepsilon$$

$$(2) \text{ AREA}_t = \beta_0 + \beta_1 \text{ AREA}_{t+5} + \varepsilon$$

In the interest of avoiding overfitting, the model was estimated using data from only 1990–2013. Using multiple models allows us to increase our country coverage as some country areas have only recently come online and did not have a full 10 years available. Table 3 presents the 10 regression models estimated.

Next, we used the coefficient results from these autoregressions to back cast the Area EFW ratings, which were then averaged to create a summary economic freedom index (EFWHAT) back to 1950. Thus under Model (1), the estimated rating for 1980 is based on the actual EFW ratings in 1985 and 1990; the 1975 estimated rating was based on the estimated 1980 rating and actual 1985 rating; the 1970 estimated rating was based on the estimated 1975 and 1980 ratings, and so on back to 1950. Table 4 shows these estimated index (EFWHAT) ratings for all countries for 1950–65.

TABLE 3
AUTOREGRESSIVE BACK CASTING MODELS ESTIMATED

	Model (1)			Model (2)	
	EFW _{t+5}	EFW _{t+10}	Constant	EFW _{t+5}	Constant
Area 1	0.736	0.08	0.841	0.793	1.202
Area 2	0.761	0.183	0.175	0.958	0.149
Area 3	0.882	−0.049	0.967	0.884	0.654
Area 4	0.975	0.073	−0.538	1.025	0.932
Area 5	0.754	0.147	0.164	0.932	0.164

Note: The five areas of the EFW index are: (1) Size of Government; (2) Legal Structure and Property Rights; (3) Sound Money; (4) Freedom to Trade Internationally; and (5) Regulation of Credit, Labor, and Business.

TABLE 4
 AUTOREGRESSIVELY BACK-CASTED ECONOMIC FREEDOM
 INDEX (EFWHAT), 1950–65

Country	1950	1955	1960	1965
Algeria	3.00	3.14	3.26	3.38
Argentina	3.05	3.17	3.27	3.37
Australia	5.66	5.83	6.01	6.20
Austria	5.30	5.44	5.58	5.74
Bahamas	4.44	4.65	4.88	5.12
Bahrain	5.27	5.43	5.62	5.81
Bangladesh	2.52	2.68	2.85	3.01
Barbados	4.61	4.79	4.97	5.16
Belgium	6.11	6.24	6.38	6.53
Benin	3.84	3.96	4.08	4.19
Bolivia	3.56	3.61	3.64	3.66
Botswana	4.75	4.87	5.00	5.12
Brazil	2.91	3.02	3.11	3.19
Bulgaria	3.98	4.18	4.39	4.61
Burundi	3.35	3.51	3.66	3.82
Cameroon	4.48	4.63	4.78	4.93
Canada	6.19	6.37	6.56	6.76
Central African Republic	3.79	3.90	4.00	4.10
Chad	4.28	4.32	4.36	4.40
Chile	4.40	4.56	4.73	4.90
China	3.91	4.03	4.14	4.26
Colombia	3.86	4.01	4.17	4.33
Congo, Dem. Rep.	3.62	3.67	3.71	3.73
Congo, Rep.	3.75	3.85	3.95	4.05
Costa Rica	4.20	4.36	4.51	4.67
Côte d'Ivoire	4.26	4.41	4.57	4.73
Cyprus	4.41	4.56	4.70	4.86
Denmark	5.52	5.63	5.74	5.86
Dominican Republic	4.33	4.47	4.61	4.75
Ecuador	3.51	3.67	3.85	4.02
Egypt	3.14	3.34	3.56	3.77
El Salvador	3.22	3.36	3.51	3.67
Fiji	4.69	4.84	5.00	5.16
Finland	5.41	5.56	5.73	5.90
France	5.11	5.23	5.36	5.48

(Continued)

TABLE 4 (*Continued*)
 AUTOREGRESSIVELY BACK-CASTED ECONOMIC FREEDOM
 INDEX (EFWHAT), 1950–65

Country	1950	1955	1960	1965
Gabon	4.25	4.39	4.54	4.70
Germany	6.05	6.18	6.32	6.47
Ghana	2.64	2.76	2.88	2.98
Greece	4.32	4.44	4.55	4.67
Guatemala	3.48	3.66	3.86	4.07
Guinea-Bissau	2.49	2.57	2.62	2.65
Guyana	2.96	3.06	3.15	3.23
Haiti	3.28	3.48	3.68	3.90
Honduras	4.53	4.64	4.72	4.91
Hong Kong	6.38	6.57	6.78	7.01
Hungary	3.60	3.74	3.88	4.02
Iceland	4.71	4.82	4.93	5.04
India	3.57	3.73	3.90	4.07
Indonesia	4.43	4.59	4.77	4.96
Iran	2.68	2.83	2.98	3.14
Ireland	5.43	5.57	5.71	5.86
Israel	4.22	4.24	4.26	4.25
Italy	5.21	5.29	5.37	5.46
Jamaica	4.18	4.30	4.43	4.55
Japan	5.86	6.04	6.23	6.43
Jordan	4.50	4.64	4.79	4.94
Kenya	4.21	4.35	4.51	4.66
Korea, South	4.44	4.58	4.73	4.88
Kuwait	4.83	4.93	5.04	5.15
Luxembourg	6.49	6.67	6.87	7.10
Madagascar	3.45	3.58	3.71	3.84
Malawi	3.87	4.02	4.16	4.32
Malaysia	5.36	5.52	5.69	5.88
Mali	3.94	4.07	4.20	4.33
Malta	4.37	4.49	4.61	4.74
Mauritius	4.47	4.92	5.12	5.33
Mexico	4.03	4.14	4.26	4.37
Morocco	4.19	4.32	4.45	4.58
Myanmar	2.04	2.25	2.46	2.67
Nepal	3.99	4.16	4.32	4.49
Netherlands	5.99	6.13	6.29	6.46

(Continued)

TABLE 4 (*Continued*)
 AUTOREGRESSIVELY BACK-CASTED ECONOMIC FREEDOM
 INDEX (EFWHAT), 1950–65

Country	1950	1955	1960	1965
New Zealand	5.41	5.53	5.66	5.79
Nicaragua	2.25	2.30	2.33	2.34
Niger	3.96	4.08	4.20	4.33
Nigeria	3.00	3.16	3.32	3.48
Norway	5.55	5.66	5.76	5.88
Oman	5.29	5.45	5.63	5.83
Pakistan	3.74	3.88	4.03	4.18
Panama	5.33	5.47	5.62	5.79
Papua New Guinea	4.69	4.85	5.01	5.19
Paraguay	3.94	4.13	4.34	4.56
Peru	2.43	2.51	2.56	2.60
Philippines	4.07	4.19	4.32	4.44
Poland	3.09	3.21	3.31	3.41
Portugal	4.84	4.94	5.03	5.13
Romania	4.44	4.53	4.64	4.77
Russia	3.89	3.86	3.82	3.77
Rwanda	4.74	4.83	4.92	5.03
Senegal	3.92	4.06	4.20	4.35
Sierra Leone	3.01	3.14	3.27	3.39
Singapore	6.23	6.38	6.54	6.71
South Africa	4.35	4.48	4.62	4.77
Spain	5.25	5.36	5.48	5.60
Sri Lanka	3.60	3.74	3.89	4.05
Sweden	5.54	5.65	5.78	5.90
Switzerland	6.38	6.55	6.74	6.96
Syria	2.69	2.79	2.88	2.96
Taiwan	5.15	5.31	5.48	5.67
Tanzania	3.09	3.22	3.35	3.47
Thailand	4.63	4.79	4.96	5.13
Togo	4.09	4.24	4.38	4.54
Trinidad & Tobago	3.77	3.93	4.09	4.35
Tunisia	3.62	3.75	3.88	4.01
Turkey	4.21	4.30	4.39	4.47
Uganda	2.87	2.93	2.97	2.99
United Arab Emirates	5.51	5.65	5.81	5.98
United Kingdom	6.25	6.40	6.57	6.75

(Continued)

TABLE 4 (*Continued*)
 AUTOREGRESSIVELY BACK-CASTED ECONOMIC FREEDOM
 INDEX (EFWHAT), 1950–65

Country	1950	1955	1960	1965
United States	6.36	6.55	6.76	6.99
Uruguay	5.34	5.44	5.55	5.66
Venezuela	4.43	4.62	4.81	5.03
Zambia	3.21	3.30	3.38	3.45
Zimbabwe	3.63	3.77	3.92	4.07

Combining the EFWNEW and EFWHAT Data to Create Economic Freedom Ratings for 1950–65

At this point we have two different sets of economic freedom ratings for 1950–80. One version, EFWNEW, is based on a set of just eight components while the second version, EFWHAT, is based on autoregressively back casted data from the regularly published EFW index. The correlation coefficient between EFWNEW and EFWHAT is 0.58, which indicates that the eight-variable index correlates reasonably well with the back casted index. The simplest way to combine the two would be to take a simple average, but we want the ultimate index to be as comparable as possible with the regular EFW index. Thus, we estimated a final model:

$$(3) \quad EFW_t = \beta_0 + \beta_1 EFWNEW_t + \beta_2 EFWHAT_t + \varepsilon$$

The time period, t , is years of overlapping data, 1970, 1975, and 1980, and the estimated equation is: $EFW_t = 1.389 + 0.22 \times EFWNEW_t + 0.56 \times EFWHAT_t$, with an Adj $R^2 = 0.67$. Both coefficients were statistically significant. The estimated coefficients from this model were used to combine the EFWNEW and EFWHAT ratings into a single combined EFWFINAL rating for 1950, 1955, 1960, and 1965. For instance, in the United States in 1950, EFWNEW had a value of 8.94 and EFWHAT 6.36, yielding a value of 6.90 for EFWFINAL. In a very different institutional context, China in 1960 had a value of 2.31 for EFWNEW and 4.14 for EFWHAT, yielding a value of 4.20.

At the end of the day, this approach uses information from later periods along with contemporaneous data to generate economic freedom ratings for the 1950–65 period. Table 5 shows this final combined EFWFINAL index for all the countries and years we have compiled.

Discussion

The top 10 and bottom 10 countries as scored by EFWFINAL in 1950 are listed in Table 6. All 10 are Western or Northern European countries and their offshoots (United States, Canada, and Australia). Among countries immediately outside the top 10 are Malaysia (11th) and Panama (14th). The lowest ranking European country as of 1950 is Greece (31st). Taiwan stands out in ranking (43rd), its score being driven by a very low score in EFWNEW. Neither Singapore nor Hong Kong, the longtime leaders in contemporary EFW, is scored for 1950, and due to data limitations Hong Kong is never scored by EFWFINAL.

The United States falls in the rankings in 1955 as data on currency restrictions and capital controls come online, falling behind Switzerland and Canada. There is sufficient data to score Japan starting that year, and it is ranked ninth among 66 countries. China also appears for the first time and ranks 56th (and a very low score of 4.25). Just ahead of China while first appearing is South Korea (50th). Ghana (66th), Guyana (63rd), and Nigeria (58th) all appear for the first time and all fall in the bottom 10 for 1955.

Luxembourg moves into the top spot in 1960, with the United States remaining in third place. Singapore appears for the first time, ranking fifth out of 83 countries, held down somewhat by a middling Judicial Independence score. Mauritius, which in the future would make the top 10 of all countries in EFW, appears first in 1960 and ranks 28th. Most countries added for 1960 appear toward the bottom of the rankings. Among those that are also in the bottom 10 are Central African Republic (82nd), Madagascar (81st), Benin (80th), and Tunisia (79th).

Another nine countries are added for 1965, all of which are in Africa. They are Rwanda (62nd), Zimbabwe (71st), Malawi (73rd), Burundi (77th), Democratic Republic of the Congo (80th), Uganda (87th), Zambia (89th), Algeria (93rd), and Tanzania (94th). Countries that had appeared in the top 10 in 1950 and had fallen out of it as of 1965 were Sweden (16th) and Norway (17th). Among the

TABLE 5
COMBINED ECONOMIC FREEDOM
INDEX (EFWFINAL), 1950-65

Country	1950	1955	1960	1965
Algeria				3.65
Argentina	3.99	4.09	4.33	4.31
Australia	6.06	6.03	6.32	6.69
Austria	5.63	5.67	6.21	6.30
Belgium	6.17	6.33	6.78	6.82
Benin			4.10	4.14
Bolivia	4.23	4.25	4.65	4.86
Brazil	4.05	4.03	4.29	4.23
Burundi				4.28
Cameroon			4.66	5.42
Canada	6.47	6.83	6.99	7.07
Central African Republic			3.96	3.94
Chad			4.18	4.19
Chile	4.72	4.84	5.24	5.35
China		4.25	4.20	4.35
Colombia	4.63	4.54	5.18	5.24
Congo, Dem. Rep.				4.23
Congo, Rep.			4.30	4.20
Costa Rica	5.16	5.59	5.67	5.50
Côte d'Ivoire			4.54	4.88
Cyprus		5.55	5.44	5.32
Denmark	5.79	5.79	6.24	6.20
Dominican Republic	4.61	4.99	5.02	4.94
Ecuador	4.59	4.75	4.89	4.98
Egypt	4.29	4.02	4.08	4.12
El Salvador	4.67	4.67	4.82	4.48
Finland	5.73	5.76	6.12	6.35
France	5.60	5.48	5.53	5.63
Gabon			4.48	4.92
Germany	6.05	6.00	6.61	6.72
Ghana		3.65	3.87	3.62
Greece	4.98	5.09	5.31	5.28
Guatemala	4.72	4.82	5.02	4.75
Guyana		3.70	3.85	3.93
Haiti		4.30	4.47	4.60

(Continued)

TABLE 5 (*Continued*)
 COMBINED ECONOMIC FREEDOM
 INDEX (EFWFINAL), 1950–65

Country	1950	1955	1960	1965
Honduras	5.58	5.49	5.55	5.49
Iceland		5.88	5.86	5.65
India	4.75	4.57	4.85	4.82
Indonesia			4.87	5.09
Iran	3.91	3.91	3.96	4.13
Ireland	5.81	5.74	5.84	5.97
Israel	4.79	4.67	4.93	4.91
Italy	5.88	5.58	5.84	5.96
Jamaica		5.27	5.30	5.91
Japan		6.01	6.22	6.45
Jordan		4.28	4.35	4.83
Kenya			5.43	5.25
Korea, South	4.47	4.71	4.99	
Luxembourg	6.46	6.63	7.26	7.37
Madagascar			4.08	4.45
Malawi				4.35
Malaysia	5.95	5.72	5.98	5.96
Mali			4.31	4.25
Malta			5.13	5.24
Mauritius			5.55	5.95
Mexico	4.94	5.09	5.16	5.33
Morocco	4.94	4.94	5.38	4.81
Myanmar	3.75	3.68	3.63	3.48
Nepal			4.40	5.15
Netherlands	5.88	6.04	6.38	6.52
New Zealand	5.85	5.70	5.80	5.87
Nicaragua	3.79	3.69	3.81	3.50
Niger			4.34	4.80
Nigeria		4.16	4.51	4.26
Norway	6.07	5.78	6.08	6.10
Pakistan	4.31	4.33	4.38	4.61
Panama	5.86	6.01	6.12	6.21
Paraguay	3.80	4.49	4.92	5.08
Peru	4.06	4.30	4.43	4.37
Philippines	5.02	4.84	4.85	5.28

(*Continued*)

TABLE 5 (Continued)
 COMBINED ECONOMIC FREEDOM
 INDEX (EFWFINAL), 1950–65

Country	1950	1955	1960	1965
Portugal	5.61	5.60	5.46	5.26
Romania			4.48	4.62
Rwanda				4.71
Senegal			4.30	4.43
Sierra Leone			4.23	4.34
Singapore			6.87	6.87
South Africa	5.01	5.27	5.35	5.39
Spain	5.28	5.27	5.59	5.72
Sri Lanka	5.15	4.72	4.51	4.56
Sweden	6.05	5.86	6.14	6.18
Switzerland	6.84	6.97	7.10	7.18
Syria	3.72	4.16	4.42	4.11
Taiwan	4.36	4.71	5.19	5.70
Tanzania				3.63
Thailand	5.16	5.28	5.28	5.38
Togo			4.59	4.68
Trinidad & Tobago		4.97	5.08	5.18
Tunisia			4.15	4.02
Turkey	4.86	4.82	4.87	5.02
Uganda				4.11
United Kingdom	6.09	6.36	6.54	6.64
United States	6.90	6.66	7.07	7.20
Uruguay	5.47	5.81	5.76	5.68
Venezuela	5.14	5.18	5.35	5.56
Zambia				4.07
Zimbabwe				4.43

East Asian Tigers, the data show Taiwan as improving, up to 26th by 1965, but not yet with South Korea, ranked 49th in 1965.

While much of the time series variation in EFWFINAL is driven by EFWHAT, we can focus on the largest movements in EFWNEW from 1950 to 1965 instead. Those that moved up most quickly by EFWNEW are, in order, Taiwan, Paraguay, Luxembourg, Bolivia, and (West) Germany. Those that had fallen the most are, in order,

TABLE 6
TOP AND BOTTOM 10 COUNTRIES AS SCORED BY COMBINED
ECONOMIC FREEDOM INDEX: 1950

Top 10 Countries	Score	Bottom 10 Countries	Score
1. United States	6.90	45. Egypt	4.29
2. Switzerland	6.84	46. Bolivia	4.23
3. Canada	6.47	47. Peru	4.06
4. Luxembourg	6.46	48. Brazil	4.05
5. Belgium	6.17	49. Argentina	3.99
6. United Kingdom	6.09	50. Iran	3.91
7. Norway	6.07	51. Paraguay	3.80
8. Australia	6.06	52. Nicaragua	3.79
9. Germany	6.05	53. Myanmar	3.75
10. Sweden	6.05	54. Syria	3.72

Sri Lanka, Myanmar, Egypt, El Salvador, and Nicaragua. One driver of this variation is the Sachs-Warner data. Of the five improving countries, all except Paraguay did so in part by becoming one of the open economies. Sri Lanka, El Salvador, and Nicaragua fell in part due to becoming one of the closed economies.

Overall, this systematic scoring and ranking reveals a few surprises while conforming sufficiently to the stylized facts of history to be credible. Taiwan and South Korea, regardless of their political alliances, had a very, very low institutional starting point at midcentury. Second, it is Luxembourg, not Singapore, that appears to be the city-state most championing liberal economic policy during this period, setting aside Hong Kong, due to the data limitations.

Conclusion

This article has presented a new economic freedom index for many countries for the 1950–65 period using a small set of eight variables (EFWNEW). This index in turn was combined with a back casted economic freedom index (EFWHAT) based on the regularly published EFW index to create an economic freedom index (EFWFINAL) for these early years. The availability of this index should improve our ability to conduct panel studies over longer periods of time.

References

- Acemoglu, D.; Johnson, S.; and Robinson, J. A. (2001) "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review* 91 (5): 1369–1401.
- Berg, C. (2015) "Classical Liberalism in Australian Economics." *Econ Journal Watch* 12 (2): 192–220.
- Burkova, L., and Lawson R. A. (2014) *Georgia's Rose Revolution: How One Country Beat the Odds, Transformed Its Economy, and Provided a Role Model for Reformers Everywhere*. Guatemala: The Antigua Forum.
- Choi, Y. B., and Yoon Y. J. (2016) "Liberalism in Korea." *Econ Journal Watch* 13 (1): 100–28.
- Chor, D. (2005) "Institutions, Wages, and Inequality: The Case of Europe and Its Periphery (1500–1899)." *Explorations in Economic History* 42 (4): 547–66.
- Clemens, M. A., and Williamson, J. G. (2004) "Why Did the Tariff-Growth Correlation Change after 1950?" *Journal of Economic Growth* 9 (1): 5–46.
- De Haan, J.; Lundstrom, S.; and Sturm J.-E. (2006) "Market-Oriented Institutions and Policies and Economic Growth: A Critical Survey." *Journal of Economic Surveys* 20 (2): 157–91.
- De La Escosura, L. P. (2016) "Economic Freedom in the Long Run: Evidence from OECD Countries (1850–2007)." *Economic History Review* 69 (2): 435–68.
- Faria, H. J., and Filardo, L. (2015) "Venezuela: Without Liberals, There Is No Liberalism." *Econ Journal Watch* 12 (3): 375–99.
- Fradejas, F. H. (2015) "Liberal Economics in Spain." *Econ Journal Watch* 12 (2): 221–32.
- Geddes, R. R., and Tennyson, S. (2013) "Passage of Married Women's Property Acts and Earnings Act in the United States." *Research in Economic History* 29: 145–89.
- Glaeser, E., and Shleifer, A. (2002) "Legal Origins" *Quarterly Journal of Economics* 117 (4): 1193–1229.
- Gwartney, J.; Block, W.; and Lawson, R. A. (1996) *Economic Freedom of the World: 1975–1995*. Vancouver: The Fraser Institute.
- Gwartney, J.; Lawson, R. A.; and Hall, J. (2016) *Economic Freedom of the World: 2016 Annual Report*. Vancouver: The Fraser Institute.

- Hall, J., and Lawson, R.A. (2014) “Economic Freedom of the World: An Accounting of the Literature.” *Contemporary Economic Policy* 32 (1): 1–19.
- Higgs, R. (1987) *Crisis and Leviathan: Critical Episodes in the Growth of American Government*. New York: Oxford University Press.
- Holcombe, R. G. (2002) *From Liberty to Democracy: The Transformation of American Government*. Ann Arbor: University of Michigan Press.
- Hough, J. F., and Grier, R. (2015) *The Long Process of Development: Building Markets and States in Pre-Industrial England, Spain, and Their Colonies*. New York: Cambridge University Press.
- Johnson, N. D., and Koyama, M. (2017) “States and Economic Growth: Capacity and Constraints.” *Explorations in Economic History* 64: 1–20.
- Kuchar, P. (2016) “Liberalism in Mexican Economic Thought, Past and Present.” *Econ Journal Watch* 13 (1): 129–67.
- Kurrild-Klitgaard, P. (2015) “Classical Liberalism and Modern Political Economy in Denmark.” *Econ Journal Watch* 12 (3): 400–31.
- La Porta, R.; Lopez-De-Silanes, F.; Shleifer, A.; and Vishny, R. W. (1997) “Legal Determinants of External Finance.” *Journal of Finance* 52 (3): 1131–50.
- La Porta, R.; Lopez-De-Silanes, F.; and Shleifer, A. (2008) “The Economic Consequences of Legal Origins.” *Journal of Economic Literature* 46 (2): 285–332.
- Lemke, J. (2016) “Interjurisdictional Competition and Married Women’s Property Acts.” *Public Choice* 166 (3): 291–313.
- Linzer, D. A., and Staton, J. K. (2015) “A Global Measure of Judicial Independence, 1948–2012.” *Journal of Law and Courts* 3 (2): 223–56.
- Manish, G. P.; Rajagopalan, S.; Sutter, D.; and White, L. H. (2015) “Liberalism in India.” *Econ Journal Watch* 12 (3): 432–59.
- Mardini, P. (2015) “The Endangered Classical Tradition in Lebanon: A General Description and Survey Results.” *Econ Journal Watch* 12 (2): 242–59.
- Marroquin, A., and Thomas, F. (2015) “Classical Liberalism in Guatemala.” *Econ Journal Watch* 12 (3): 460–78.
- Mateusz, M. (2015) “Liberal Economics in Poland.” *Econ Journal Watch* 12 (2): 233–41.

- Murphy, R. H., and Stansel, D. (2017) "Subnational Economic Freedom of the U.S. States in the Nineteenth Century." O'Neil Center for Global Markets and Freedom Working Paper, Southern Methodist University.
- Prokopjević, M., and Tasić, S. (2015) "Classical Liberal Economics in Ex-Yugoslav Nations." *Econ Journal Watch* 12 (2): 260–73.
- Sachs, J. D., and Warner, A. (1995) "Economic Reform and the Process of Global Integration." *Brookings Papers on Economic Activity* 26: 1–118.
- Sima, J., and Nikodym, T. (2015) "Classical Liberalism in Czech Republic." *Econ Journal Watch* 12 (2): 274–92.
- Summers, R., and Heston, A. (1984) "Improved International Comparisons of Real Product and Its Composition: 1950–1980." *Review of Income and Wealth* 30 (2): 207–19.
- Wang, N., and Coase, R. (2012) *How China Became Capitalist*. New York: Palgrave Macmillan.
- Yergin, D., and Stanislaw, J. (1998) *The Commanding Heights: The Battle between Government and the Marketplace That Is Remaking the Modern World*. New York: Simon & Schuster.

FORMER CENTRALLY PLANNED ECONOMIES 25 YEARS AFTER THE FALL OF COMMUNISM

James D. Gwartney and Hugo M. Montesinos

A little more than a quarter of a century has passed since the collapse of communism, which makes this an ideal time to evaluate the subsequent development of once communist nations. Which countries have moved the most toward economic liberalization? How have the former centrally planned (FCP) economies performed in recent decades? How have their political institutions evolved during the transition era and beyond? What lessons can be learned from the experience of these economies? This article will address each of those questions.

In some ways, the experience of the FCP countries constitutes a natural economic experiment. There is considerable diversity in the paths they have followed. Some moved rapidly toward economic reform and liberalization following the collapse of communism, but others moved more slowly, and still others have undertaken little or no reform. Some of the FCP countries had relatively high per capita incomes prior to the fall of communism, while others were exceedingly poor. Some experienced lengthy and painful transitions, while

Cato Journal, Vol. 38, No. 1 (Winter 2018). Copyright © Cato Institute. All rights reserved.

James D. Gwartney is a Professor of Economics and the Gus A. Stavros Eminent Scholar at Florida State University. Hugo M. Montesinos is a doctoral student and resident econometrician at Florida State University. This report was initially prepared for presentation at the annual meeting of the Economic Freedom of the World network held September 28–30, 2017, in Kiev, Ukraine. The full report containing the tables with detailed data referred to in this article is available at <http://coss.fsu.edu/hilton/node/6>.

others made the move from central planning to markets more smoothly. Some of these countries are now highly democratic, while others are still governed by authoritarian political regimes. As we examine the experience of the FCP economies, we will do so with an eye to what can be learned about institutions, economic growth, and the development process.

The article is organized in the following manner. Section 1 examines the path of economic liberalization of 25 FCP economies from 1995 to 2015. Section 2 presents data on various indicators of economic performance during this same time frame. Section 3 focuses on the evolution of political institutions (e.g., protection of civil liberties, democracy, and control of corruption) in the FCP countries. Section 4 compares the income levels and growth rates of these economies relative to the world's high-income countries and other developing economies. Section 5 examines the determinants of economic growth and life satisfaction for 122 countries and considers the implications for the FCP economies. Section 6 analyzes areas where the FCP economies have made substantial moves toward economic liberalization, as well as a major deficiency—low-quality legal systems—that is likely to restrain their future progress. The concluding section summarizes and considers the implications of the analysis.

Economic Liberalization and the Former Centrally Planned Economies

The Economic Freedom of the World (EFW) project provides a measure of the degree to which the institutions and policies of various countries are consistent with economic freedom (Gwartney, Lawson, and Hall 2017). This measure uses more than 40 different variables to construct a summary index of economic freedom. The EFW index now covers 159 countries and the data are available for 123 countries since 1995. This data set makes it possible to identify cross-country differences in economic freedom and to track changes across time.

The EFW index is designed to measure the degree to which the institutions of a country are supportive of (1) personal choice, (2) voluntary exchange, (3) open entry into markets, and (4) protection of individuals and their property from aggression by others. Because economic freedom facilitates and encourages gains from trade, entrepreneurship, innovation, and capital formation, economic

theory indicates that it is an important source of economic growth and development. Several empirical studies have found that this is indeed the case.¹ Moreover, economic freedom permits individuals to mold and shape their lives according to their preferences. Over and above the impact on income, this may enhance quality of life.² There are 25 former centrally planned economies for which the EFW data are now available. These data are available continuously throughout the 1995–2015 period for 14 of these countries.

Table 1 provides the EFW summary ratings and worldwide rankings (in parentheses) for these 25 countries (when available) for 1995, 2000, 2005, 2010, and 2015. Seven of the FCP economies (Georgia, Estonia, Lithuania, Latvia, Romania, Armenia, and Albania) had a 2015 EFW summary rating of 7.5 or higher. Worldwide, these seven countries all ranked in the top quartile among the 159 countries for which the EFW data were available. Moreover, these countries have achieved dramatic increases in economic freedom. While the Baltic states all ranked in the Top 20 in 2015, in 1995 Estonia was 57th, Lithuania 80th, and Latvia 75th. Romania ranked 20th in 2015, but it was a late reformer. Romania's worldwide ranking was 118th in 1995 and 107th in 2000 (among the 123 countries included in the index during those years). Albania has steadily improved both its rating and ranking, moving up from 96th in 1995 to 63rd in 2005 and 32nd in 2015. While the EFW data were unavailable for Georgia and Armenia in 1995 and 2000, the ratings and rankings of both have increased since 2005.

A group of nine other countries (Czech Republic, Bulgaria, Poland, Slovak Republic, Hungary, Kazakhstan, Macedonia, Croatia, and Slovenia) had 2015 EFW summary ratings between 7.0 and 7.5. Worldwide, the 2015 rankings of these countries ranged from 42nd for the Czech Republic to 73rd for Slovenia. Thus, each of these nine countries ranked in the second quartile among the 159 countries included in the EFW data set in 2015. These nine countries make up the middle group in terms of economic liberalization among the 25 FCP economies.

¹See for example Berggren (2003), Dawson (1998, 2003), Faria and Montesinos (2009), Feldmann (2017), and Nystrom (2008).

²For analysis of this topic, see Bjørnskov, Dreher, and Fischer (2010); Pitlik and Rode (2016); and Rode (2013).

TABLE 1
ECONOMIC FREEDOM RATINGS AND RANKINGS FOR THE
25 FORMER CENTRALLY PLANNED COUNTRIES, 1995–2015

Country	1995	2000	2005	2010	2015
<i>Top EFW Group: 2015 EFW \geq 7.50</i>					
Georgia			7.42 (33)	7.50 (27)	8.01 (8)
Estonia	6.12 (57)	7.48 (23)	7.96 (11)	7.82 (10)	7.95 (10)
Lithuania	5.51 (80)	6.90 (53)	7.37 (40)	7.47 (29)	7.92 (13)
Latvia	5.59 (75)	7.13 (39)	7.42 (33)	7.23 (50)	7.75 (17)
Romania	3.83 (118)	5.37 (107)	7.24 (49)	7.30 (45)	7.72 (20)
Armenia			7.31 (44)	7.56 (24)	7.60 (29)
Albania	5.10 (96)	6.20 (73)	6.96 (63)	7.35 (37)	7.54 (32)
<i>Middle EFW Group: 2015 EFW between 7.00 and 7.50</i>					
Czech Rep.	5.99 (72)	6.71 (62)	6.98 (62)	7.22 (52)	7.46 (42)
Bulgaria	4.8 (101)	5.52 (104)	6.95 (64)	7.30 (45)	7.39 (48)
Poland	5.28 (90)	6.58 (72)	6.89 (67)	7.12 (61)	7.34 (51)
Slovak Rep.	5.25 (83)	6.85 (57)	7.63 (20)	7.47 (29)	7.31 (53)
Hungary	6.15 (58)	7.03 (47)	7.20 (52)	7.31 (44)	7.30 (54)
Kazakhstan			6.83 (69)	6.94 (71)	7.18 (66)
Macedonia			6.36 (86)	6.93 (72)	7.17 (67)
Croatia	4.98 (94)	6.12 (78)	6.47 (83)	6.68 (88)	7.02 (72)
Slovenia	5.22 (87)	6.63 (71)	6.91 (66)	6.82 (80)	7.00 (73)
<i>Bottom EFW Group: 2015 EFW < 7.00</i>					
Kyrgyz Rep.			6.61 (79)	6.61 (94)	6.89 (80)
Tajikistan				6.28 (113)	6.80 (82)
Montenegro			6.35 (93)	7.33 (40)	6.77 (85)
Serbia			5.96 (109)	6.56 (97)	6.75 (88)
Bosnia & Hrgvna			6.18 (100)	6.63 (91)	6.61 (99)
Russia	4.48 (107)	5.39 (106)	6.24 (98)	6.54 (98)	6.60 (100)
Moldova			6.67 (73)	6.58 (96)	6.56 (102)
Azerbaijan			6.04 (106)	5.97 (127)	6.38 (114)
Ukraine	3.39 (123)	4.69 (117)	5.81 (118)	5.90 (133)	5.38 (149)
Number of Countries Included in the Index	123	123	141	153	159

NOTE: The table is sorted according to the 2015 EFW summary rating. The worldwide EFW ranking, each year, is in parentheses. The total number of countries included in the worldwide EFW ranking is in the last row of the table.

SOURCE: Gwartney, Lawson, and Hall (2017).

The Czech Republic is the highest ranked country in the middle group, and it has shown significant improvement. It ranked 42nd in 2015, up from 72nd in 1995. Other countries in this group have registered even more impressive gains in economic freedom. For example, Bulgaria's 2015 worldwide ranking was 48th, up from 101st in 1995 and 104th in 2000. Poland ranked 51st in 2015, up from 90th in 1995 and 72nd in 2000. The ranking of the Slovak Republic rose from 83rd in 1995 to 20th in 2005, but it has subsequently receded to 53rd in 2015. The movements toward economic freedom of Hungary, Croatia, and Slovenia during 1995–2015 were more modest.

Finally, there is another set of nine FCP economies with 2015 EFW summary ratings of less than 7.0. This set of countries is composed of the Kyrgyz Republic, Tajikistan, Montenegro, Serbia, Bosnia and Herzegovina, Russia, Moldova, Azerbaijan, and Ukraine. The worldwide rankings in 2015 of these countries ranged from 80th for the Kyrgyz Republic to 149th for Ukraine. Except for Ukraine, the 2015 ranking for each of these countries placed them in the third quartile worldwide. Ukraine was in the fourth quartile. In 2015, these nine countries were the least economically free among the FCP economies. Further, there is little evidence of improvement among the countries in this group. These countries ranked in the bottom half worldwide during 1995–2005, and this was still true in 2015. The case of Russia is typical. Russia ranked 107th in 1995, 98th in 2005, and 100th in 2015.

Indicators of Economic Performance: 1995–2015

How does the performance of the FCP economies that have made more substantial moves toward economic freedom compare with those that have been slow to liberalize? In order to provide insight on this question, this section will examine the income levels, growth rates, international trade sectors, and foreign investment of the FCP economies during 1995–2015.

Per Capita GDP and Growth

Table 2 shows the per capita GDP (2011 PPP dollars) for the high, middle, and low economic freedom FCP groups in the years 1995 (column 1) and 2015 (column 2). Within the most economically free group, the countries with the highest per capita 2015 GDP were Estonia, Lithuania, Latvia, and Romania. The 2015 per capita GDP

TABLE 2
1995 AND 2015 PER CAPITA GDP AND ANNUAL GROWTH
RATE OF REAL PER CAPITA GDP FOR THE FORMER
CENTRALLY PLANNED ECONOMIES, 1995–2015

Country (2015 EFW Rank)	Per Capita GDP (2011 PPP dollars)		Annual Growth Rate (percent)		
	1995	2015	1995–2015	2000–2015	2005–2015
<i>Top EFW Group: 2015 EFW ≥ 7.50</i>					
Georgia (8)	2,295	9,025	7.09	7.02	6.29
Estonia (10)	11,362	27,329	4.49	3.76	1.83
Lithuania (13)	9,357	26,971	5.44	5.44	3.83
Latvia (17)	8,272	23,057	5.26	4.96	2.8
Romania (20)	10,546	20,538	3.39	4.56	3.43
Armenia (29)	2,173	8,180	6.85	7.1	4.32
Albania (32)	4,129	11,025	5.03	4.78	3.98
Simple Mean	6,876	18,018	5.36	5.37	3.78
Pop. Wtd. Mean	8,202	18,349	4.54	5.13	3.81
<i>Middle EFW Group: 2015 EFW between 7.00 and 7.50</i>					
Czech Rep. (42)	19,215	30,381	2.32	2.45	1.67
Bulgaria (48)	8,446	17,000	3.56	4.36	2.97
Poland (51)	11,300	25,299	4.11	3.67	3.94
Slovak Rep. (53)	13,184	28,254	3.88	4.04	3.5
Hungary (54)	15,244	24,831	2.47	2.22	1.08
Kazakhstan (66)	8,283	23,522	5.36	5.9	3.92
Macedonia (67)	7,641	12,760	2.6	2.65	3.12
Croatia (72)	12,625	20,636	2.49	1.82	0.54
Slovenia (73)	18,431	29,097	2.31	1.66	0.77
Simple Mean	12,708	23,531	3.23	3.20	2.39
Pop. Wtd. Mean	12,044	24,646	3.78	3.72	3.07
<i>Bottom EFW Group: 2015 EFW < 7.00</i>					
Kyrgyz Rep. (80)	1,696	3,238	3.28	3.01	3.17
Tajikistan (82)	1,270	2,641	3.73	5.52	4.46
Montenegro (85)	10,205	15,291	2.27	2.82	2.98
Serbia (88)	7,393	13,278	2.97	3.45	1.99
Bosnia & Hrzgva (99)	1,827	10,902	9.34	3.69	2.75
Russia (100)	12,813	24,124	3.21	3.67	2.24
Moldova (102)	2,605	4,747	3.04	4.89	3.68

(Continued)

TABLE 2 (*Continued*)
 1995 AND 2015 PER CAPITA GDP AND ANNUAL GROWTH
 RATE OF REAL PER CAPITA GDP FOR THE FORMER
 CENTRALLY PLANNED ECONOMIES, 1995–2015

Country (2015 EFW Rank)	Per Capita GDP (2011 PPP dollars)		Annual Growth Rate (percent)		
	1995	2015	1995–2015	2000–2015	2005–2015
Azerbaijan (114)	3,320	16,699	8.41	9.2	7.57
Ukraine (149)	5,060	7,465	1.96	2.99	0.3
Simple Mean	4,498	10,932	4.50	4.36	3.24
Pop. Wtd. Mean	9,630	18,271	3.30	3.83	2.22

NOTE: This table is sorted according to the 2015 EFW summary rating. The worldwide 2015 EFW ranking, out of 159 countries, is in parentheses. The population weighted mean for each group was computed using the 2015 population. In the case of Montenegro, the earliest per capita GDP figure available from the World Bank is for the year 1997. Therefore, the per capita GDP for Montenegro reported in the table is for 1997 rather than 1995. Similarly, the growth figure for Montenegro is for 1997–2015 rather than 1995–2015.

SOURCE: World Bank (2017).

for each of these countries exceeded \$20,000. In the middle group, seven of the nine countries (Czech Republic, Poland, Slovak Republic, Hungary, Kazakhstan, Croatia, and Slovenia) all registered a 2015 per capita GDP of greater than \$20,000. In this group, only Bulgaria and Macedonia failed to reach that benchmark. In the group with the lowest EFW ratings in 2015, only Russia achieved a 2015 per capita GDP of greater than \$20,000. Four of the countries in this group (Kyrgyz Republic, Tajikistan, Moldova, and Ukraine) had a 2015 per capita GDP of less than \$10,000.

With regard to the per capita GDP of the three groups, the simple and population weighted mean for the middle group was the highest, followed by the most-free group. The group with the lowest EFW ratings also had the lowest 2015 mean per capita income levels.

Table 2 also presents figures for the annual real growth rate of per capita GDP of the 25 countries during 1995–2015, 2000–15, and

2005–15. As column 3 shows, six of the seven countries in the most-free group had growth rates of 4 percent or higher during 1995–2015. The exception was Romania, which did not begin to liberalize until after 2000 (see Table 1). After adopting reforms supportive of economic freedom, Romania achieved an annual growth rate of per capita GDP of 4.56 percent during 2000–15. The per capita GDP annual growth rate for five of the seven countries in the most economically free group exceeded 5 percent during 1995–2015. The simple mean and population weighted growth rates for the most-free group were 5.36 percent and 4.54 percent respectively.

Among the countries in the middle group, the annual growth rates of Poland, Bulgaria, Slovak Republic, and Kazakhstan were the most impressive. However, only Poland and Kazakhstan were able to achieve an annual growth rate greater than 4 percent during 1995–2015. The simple mean annual growth of per capita GDP was 3.23 percent for the middle group, while the population weighted mean was 3.78 percent.

The simple and population weighted means for growth during 1995–2015 of the least-free group were 4.50 percent and 3.30 percent respectively. Among the eight countries in the least-free group, only Bosnia and Herzegovina and Azerbaijan were able to achieve an annual growth rate greater than 4 percent during 1995–2015. Interestingly, special circumstances underlie the growth of both of these countries. Compared to the size of its economy, Azerbaijan is the leading oil exporter among the FCP economies. The high oil prices of 2002–14 were a major factor underlying its strong growth. The 1995 per capita GDP of Bosnia and Herzegovina was depressed by the aftermath of civil war, and therefore its 9.34 percent annual growth rate during 1995–2015 was elevated. Its real growth rates of 3.69 percent and 2.75 percent during 2000–15 and 2005–15 respectively are more indicative of its long-term growth path.

The FCP countries that liberalized the most generally grew more rapidly during 1995–2015 than their counterparts that were slow to reform. Consider the number of countries in each of the three groups that achieved an annual growth rate of at least 4 percent during the two-decade time frame. Six of the seven countries in the most economically free group achieved this benchmark, but only two of the nine countries in the middle group and only two of the eight countries in the least-free group were able to achieve this

figure. Moreover, the population weighted mean annual growth rate of the most-free group was 4.54 percent, compared to 3.78 percent for the middle group and 3.30 percent for the least-free group.

Like column 3, columns 4 and 5 of Table 2 present growth data; the difference is that the length of the periods examined are 15 years (2000–15) and 10 years (2005–15). The growth figures for 2000–15 have a similar pattern as those for the 20-year period. The population weighted mean annual growth rate of per capita GDP during 2000–15 of the most-free group was 5.13 percent, compared to 3.72 percent for the middle group, and 3.83 percent for the least-free group. Again, the mean growth rates of the middle and least-free group are elevated by the high growth rates of the oil exporting countries of Kazakhstan and Azerbaijan. Comparison of the 15- and 20-year growth figures for Romania and Bulgaria also highlights an interesting point. As Table 1 indicates, these countries were late reformers. Neither undertook significant reforms until after 2000. Note, their growth rates were higher during 2000–15 than for 1995–2015.

Comparison of the growth rates across the three periods also highlights a key point: the growth rates of the most recent decade were generally lower than for the earlier periods. Only two countries—Macedonia and Montenegro—had a higher growth rate of real per capita GDP in the most recent period, 2005–15, than during the two longer time frames. This illustrates that the growth rates of most of these economies have slowed in recent years. Nonetheless, the annual growth rate of per capita GDP during 2005–15 of 10 of the 25 FCP countries (Georgia, Lithuania, Armenia, Albania, Poland, Kazakhstan, Slovak Republic, Tajikistan, Moldova, and Azerbaijan) exceeded 3.5 percent. Thus, while growth has generally slowed, it remains relatively strong among these economies. As we proceed, the relationship between economic freedom and the growth rate of the FCP economies will be examined in more detail.

Growth of the Trade Sector

International trade promotes gains from specialization, economies of scale, and importation of innovative products and production methods. Further, international trade makes it possible for both

consumers and producers in a domestic economy to gain from greater integration into the worldwide network of markets. Thus, economic analysis indicates that trade openness and expansion in trade will elevate economic growth.

The ratio of exports plus imports divided by GDP provides a straightforward measure for the size of the trade sector. Comparison of the beginning and ending time frames provides insight on changes in the size of the trade sector over the two-decade period.

Except for Armenia, all the countries in the most economically free group experienced substantial increases in trade as a share of GDP. The mean size of the trade sector for this group rose from 79.5 percent during 1996–2000 to 111.9 percent in 2011–15, an increase of 40 percent. The countries in the middle group also experienced sizeable expansions in international trade. On average, trade as a share of GDP for this group rose from 86.4 percent during 1996–2000 to 125.0 percent during 2011–15, an increase of approximately 45 percent. Clearly the countries in both the top and middle groups in terms of economic freedom experienced substantial increases in the size of their trade sectors. However, the situation was quite different for the least economically free group. Only three of the nine countries in this group—Kyrgyz Republic, Montenegro, and Serbia—experienced significant expansions in trade. The size of the trade sector for the other six countries in this group was either similar or smaller in 2015 than in the late 1990s. The mean for this group was 93.1 percent in 2011–15, virtually unchanged from 93.7 percent in 1996–2000.

Eight FCP countries (Czech Republic, Estonia, Hungary, Latvia, Lithuania, Slovak Republic, Slovenia, and Poland) joined the European Union (EU) in 2004, and two others (Romania and Bulgaria) joined in 2007. Still later, Croatia joined the EU in 2013. In addition to its central government functions, the EU is a customs union. In fact, it is an outgrowth of a free trade agreement among several European countries. The EU sets common tariff rates and international trade policy for all member countries, but there are no tariffs or restrictions on the movement of goods within the union.

Joining the EU will generally reduce trade barriers and enhance the size of the trade sector of an FCP country. There are two reasons why this will be the case. First, joining the EU will provide domestic consumers and producers with a vastly larger “free trade” market. Thus, trade with partners in other EU countries will generally increase. Second, because tariff rates and other trade restrictions

imposed by the EU are relatively low, the trade barriers with non-EU members will also tend to decline. This will be particularly true if the trade restrictions of the joining member were high prior to membership in the union.

Did joining the EU reduce trade barriers and lead to an expansion in trade? There is evidence this was the case. All ten of the FCP countries that joined the EU during 2004–07 had substantially larger trade sectors in 2011–15 than during 1996–2000. Further, the increases in the size of their trade sectors were exceedingly large. For example, between 1996–2000 and 2011–15, international trade as a share of GDP soared in Lithuania from 88 percent to 159 percent. In the Czech Republic, the size of the trade sector rose from 87 percent to 151 percent; in the Slovak Republic, the increase was from 110 percent to 180 percent; in Poland, the parallel increase was from 53 percent to 91 percent. Similarly, between 1996–2000 and 2011–15 the trade sector of Hungary rose from 107 percent to 169 percent and that of Slovenia soared from 97 percent to 144 percent. Latvia and Bulgaria experienced similar large increases in the size of their trade sectors soon after joining the EU. Moreover, the expansions in the trade sector of the FCP countries that joined the EU were substantially greater than those achieved by the non-EU FCP countries. These trade increases are consistent with the view that joining the EU reduced trade barriers, enhanced international trade, and promoted integration into the world economy.

Foreign Direct Investment

Foreign direct investment (FDI) plays a key role in the growth process. There are several reasons why this is the case. First, almost all FDI is private. Thus, it reflects investor confidence in the institutions and future of a country. Second, FDI is an important source of innovation and technology transfers among countries. This is particularly important for developing economies because they often lag well behind their higher-income counterparts in these areas. Finally, FDI is also a source of financing for capital investments, an ingredient that is often in short supply in lower-income, developing economies.

Net FDI as a share of GDP during 1995–2015 was derived for the 25 FCP economies. FDI increased as a share of GDP in most of these economies during the first decade of this century, but it has

declined substantially since 2010. For example, the average net FDI as a share of the economy for the seven countries with the highest EFW ratings rose from 4.6 percent during 1996–2000 to 5.4 percent in 2001–05 and 7.5 percent in 2006–10, but it then receded sharply to 4.6 percent during 2011–15. Further, this pattern—elevated levels of net FDI during 2001–10, but declines during the past five years—was present for the mean values of the other two groups. The declining levels of net FDI as a share of the economy are a troubling sign. This is likely to slow the rate of future economic growth.

Economic Record of the FCP Countries

The economic record of the FCP countries during 1995–2015 was impressive. This was particularly true for the seven FCP countries that moved the most toward economic liberalization. The average growth of real per capita GDP of these seven countries exceeded 5 percent during 1995–2015. Real per capita GDP more than doubled in six of these seven countries during the two decades. The late-reforming Romania was the exception, and its per capita GDP almost doubled (it increased by 95 percent) in just 15 years following adoption of liberal reforms early in this century. While the real GDP growth of the middle group was slower, it was still impressive. Most of the countries in the most-free and middle group experienced large increases in international trade, an in-flow of FDI, and rapid growth rates. Economic growth, expansion in international trade, and FDI lagged in most of the least-free economies, but even this group achieved a population weighted annual growth of per capita GDP of 3.3 percent during 1995–2015.

Civil Liberties and Political Institutions

The FCP economies have a history of authoritarianism, political corruption, and abuse of civil liberties. Thus, sensitivity to the operation of political institutions is an issue of considerable importance.

Freedom House has provided ratings for both civil liberties and political rights annually since 1972. The Freedom House rating scale ranges from 1 (most free) to 7 (least free). Countries with a rating of 1 or 2 are classified as “free,” 3, 4, or 5 as “partly free,” and 6 or 7 as “not free.”

Freedom House classifies seven of the 25 FCP economies as free for both civil liberties and political rights throughout the entire

period. These seven countries are Estonia, Lithuania, Latvia, Czech Republic, Poland, Hungary, and Slovenia. By 2015, Romania, Bulgaria, Slovak Republic, Croatia, and Serbia joined the “free” group for both civil liberties and political rights. Except for Serbia, all of the countries with civil liberties and political rights classifications as “free” are from the two groups with the highest EFW ratings. Moreover, other than Serbia, none of the countries in the bottom EFW group was classified as “free” in both civil liberties and political rights during any of the years. Freedom House rates Tajikistan, Russia, and Azerbaijan as “not free” in both civil liberties and political rights in 2015. The ratings for Russia are particularly interesting because of their persistent deterioration. Russia’s civil liberties rating was 4 in 1995, 5 during 2000–10, and 6 in 2015. In political rights, Russia’s rating receded from 3 in 1995 to 5 in 2000, and 6 during 2005–15.

The Polity IV data set (Marshall, Gurr, and Jagers 2016) provides information on both democracy and constraints on the executive. The scale for the democracy variable ranges from minus 10 (strongly autocratic) to plus 10 (strongly democratic). The Polity IV data indicate that most of the FCP economies moved toward democracy during 1995–2015. By 2015, only three countries, Kazakhstan, Tajikistan, and Azerbaijan, were classified as autocratic (negative rating). Most of the 25 countries have positive ratings of 8 or more. In the most economically free group, only Georgia and Armenia had a 2015 rating of less than 8, and in the middle group, only Kazakhstan failed to meet this benchmark. However, in the least-free group, five countries—Kyrgyz Republic, Tajikistan, Russia, Azerbaijan, and Ukraine—had democracy ratings of less than 8.

The scale of the polity data for constraints on the executive variable ranges from 1 (no limitations on executive actions) to 7 (accountability groups such as legislatures have the power to constrain executive actions). As in the case of democracy, the ratings for constraints on the executive were higher in 2015 than was true two decades earlier. In 2015, all countries of the most-free group had ratings of 7 except for Georgia (rating of 6) and Armenia (rating of 5). In the middle group, eight of the nine countries had a rating of 7; the exception was Kazakhstan with a rating of 2. In the least-free group, four of the nine countries—Kyrgyz Republic, Montenegro, Serbia, and Moldova—had a rating of 7. However, the constraints on the executive were weak for four other countries in this group: Tajikistan

(rating of 3), Russia (rating of 4), Azerbaijan (rating of 2), and Ukraine (rating of 5). While there are countries with democratic political institutions in each of the three groups, countries in the least economically free group are more likely to be less democratic and have weaker constraints on the executive.

Transparency International (2015) provides data on corruption, which is defined as “the abuse of public office for private gain.” These data are used to develop the Corruption Perception Index (CPI). The CPI ranges from 0 (highly corrupt) to 100 (highly clean). The CPI increased for almost all of the 25 FCP economies, indicating a reduction in the level of corruption. The CPI was unavailable for a number of countries in 1995 and 2000. Thus, we will focus on the ratings during 2005–15. For the most-free group, the average CPI increased from 37.1 in 2005 to 50.7 in 2015. For the middle group, the average CPI rose from 39.8 in 2005 to 49.1 in 2015. For the least-free group, the average CPI rose from 25.3 in 2005 to 32.7 in 2015. The 2015 average CPI is considerably higher for groups 1 and 2 than for group 3. The following four countries had a 2015 CPI of 60 or higher: Estonia (70), Lithuania (61), Poland (62), and Slovenia (60). In contrast, the 2015 CPI was less than 30 for the following countries: Kazakhstan (28), Kyrgyz Republic (28), Tajikistan (26), Russia (29), Azerbaijan (29), and Ukraine (27). Note that all four of the countries with the highest 2015 CPI are from the two groups with the highest 2015 EFW ratings. In contrast, five of the six countries (Kazakhstan is the exception) with the lowest 2015 CPI are from the group with the lowest 2015 EFW rating.

Pulling the data on political institutions together, the following nine countries had 2015 political institutions most consistent with protection of civil liberties, political democracy, and absence of corruption: Estonia, Lithuania, Latvia, Czech Republic, Poland, Slovak Republic, Hungary, Croatia, and Slovenia.³ In contrast, the political institutions of Kazakhstan, Tajikistan, Russia, and Azerbaijan were most inconsistent with civil liberties protection, political democracy, and absence of corruption.

³In 2015, these countries had civil liberties and political rights ratings of 1 or 2; democracy scores of 8, 9, or 10; constraints on the executive of 6 or 7; and a CPI of 50 or more.

The Income of the FCP Economies Compared to the World's High-Income Countries and Other Developing Economies

This section will compare the relative per capita GDP of the FCP economies with the 21 high-income countries and the 82 other developing economies for which the economic freedom data were available for 1995–2015. The FCP countries with the highest income levels were Estonia, Lithuania, Czech Republic, Slovak Republic, and Slovenia. By 2015, the per capita GDP for each of these five countries had risen to 60 percent or more of the mean for the 21 high-income countries. The countries with the lowest 2015 income levels were Armenia, Kyrgyz Republic, Tajikistan, Moldova, and Ukraine. The per capita GDP of each of these five countries was less than 20 percent of the comparable mean for the high-income group.

Table 3 provides the annual growth rate of per capita GDP for the 21 countries in the high-income group, 16 high-income European countries, and for the 82 non-FCP developing economies. The per capita growth data are also provided for the 25 FCP economies according to their 2015 EFW summary rating. Both the simple and population weighted mean growth rates are provided for three different time periods—1995–2015, 2000–15, and 2005–15.

How do the growth rates of the FCP countries compare to the other groups? As Table 3 shows, the centrally planned economies grew more rapidly than the high-income countries throughout the 1995–2015 period. For example, the simple mean annual growth rate of the top, middle, and bottom groups (according to 2015 EFW ratings) were 5.36 percent, 3.23 percent, and 4.5 percent, respectively. Each of these rates was well above the simple mean of 1.5 percent for the world's 21 high-income countries and the 1.52 percent annual growth rate for the 16 high-income European countries. The population weighted mean annual growth rates for the top (most free), middle, and bottom (least free) FCP groups during 1995–2015 were 4.54 percent, 3.78 percent, and 3.3 percent, respectively. Again, these annual growth figures are all considerably higher than the 1.25 percent for the 21 high-income countries of the world and 1.16 percent for the 16 European countries. When these comparisons are also made for the 2000–15 and 2005–15 periods, the pattern of the results is the same: the growth rate for each of the FCP groups exceeds that of the high-income countries.

TABLE 3
ANNUAL GROWTH RATES OF PER CAPITA GDP OF FORMER
CENTRALLY PLANNED ECONOMIES AND OTHER SETS
OF COUNTRIES

Group of Countries	1995–2015	2000–2015	2005–2015
<i>Simple Average Annual Growth Rate (%)</i>			
21 High-income	1.50	0.94	0.59
16 High-income European	1.52	0.90	0.54
Other 82 developing	2.03	2.18	2.15
Other 82 (excluding China and India)	1.91	2.05	2.02
China and India	7.01	7.39	7.57
25 Former Centrally Planned	4.27	4.23	3.09
7 FCP—Top 2015 EFW group	5.36	5.37	3.78
9 FCP—Middle 2015 EFW group	3.23	3.20	2.39
9 FCP—Bottom 2015 EFW group	4.50	4.36	3.24
<i>Population Weighted Average Annual Growth Rate (%)</i>			
21 High-income	1.25	0.80	0.52
16 High-income European	1.16	0.67	0.40
Other 82 developing	4.75	5.14	5.21
Other 82 (excluding China and India)	2.41	2.82	2.77
China and India	7.06	7.43	7.61
25 Former Centrally Planned	3.55	3.93	2.61
7 FCP—Top 2015 EFW group	4.54	5.13	3.81
9 FCP—Middle 2015 EFW group	3.78	3.72	3.07
9 FCP—Bottom 2015 EFW group	3.30	3.83	2.22

NOTE: The 21 high-income countries are Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Iceland, Ireland, Italy, Japan, Luxembourg, the Netherlands, New Zealand, Norway, Spain, Sweden, Switzerland, the United Kingdom, and the United States. The 16 high-income European countries comprise the 21 high-income countries, minus Australia, Canada, Japan, New Zealand, and the United States. There were 123 countries with continuous EFW data from 1995 to 2015. The 21 high-income industrial countries and 14 FCP economies are included in this group. Thus, the EFW data were available for 88 developing economies. However, the per capita GDP data of six of these countries (Venezuela, Syria, Papua New Guinea, Guyana, Haiti, and Taiwan) were unavailable in the World Bank data in either 1995 or 2015. Thus, the growth rate data from the World Bank were available for 82 non-FCP developing economies

SOURCE: World Bank (2017).

Turning to a comparison between the FCP economies and the other 82 developing countries, the simple average annual growth rate of the FCP groups nearly always exceeds the simple average for the 82 developing economies. For example, the simple mean annual growth rate for 1995–2015 of the 82 developing economies was 2.03 percent, compared to the annual growth rates of 5.36 percent, 3.23 percent, and 4.50 percent for the top, middle, and bottom FCP groups. The pattern was similar for the 15- and 10-year comparisons: the simple average annual growth rates of the FCP economies were generally greater than the simple average for the 82 developing countries.

However, the pattern changes when the population weighted figures are used for the comparisons. The population weighted mean annual growth rates for the 82 developing economies are generally greater than the parallel rates for the FCP countries. For example, the population weighted mean annual growth rate for the 82 developing countries during 1995–2015 was 4.75 percent compared to 4.54 percent, 3.78 percent, and 3.30 percent for the top, middle, and bottom groups among the FCP economies. The population weighted growth rates for the 82 developing economies are driven by the high growth rates of China and India, the world's two most populous countries. When these two countries are omitted from the developing group, the mean annual growth rate of the remaining 80 countries is substantially lower. When the FCP groups are compared with the developing countries without China and India, the growth rates of the FCP economies are generally higher than those of the 80 developing economies.

To summarize: the growth rates of the FCP economies are generally higher than the growth rates of the world's 21 high-income countries, the 16 high-income European economies, and the developing economies of the world, except for China and India. This pattern holds for both the simple average and the population weighted average growth rates and for each of the three periods.

Determinants of Economic Growth and Life Satisfaction

A regression model of economic growth was developed and tested across 122 countries.⁴ The dependent variable was the annual growth

⁴The regression equations can be found in Table 15 of the full report, available at <https://coss.fsu.edu/hilton/sites/default/files/1b.%20FCP%20economies%20-%20Complete%20Study.pdf>.

rate of real per capita GDP. The analysis was conducted for three periods: 1995–2015, 2000–15, and 2005–15. The lengthier time frames, particularly the 15- and 20-year periods, will minimize the impact of business cycle factors on the measurement of long-term growth. The model included the following variables with the expected sign indicated in parentheses: 1995 per capita GDP (minus), economic freedom (plus), log of population (plus), initial percentage of female population in prime working age 25–59 group⁵ (plus), change in the percentage of this population during the period (plus), net foreign direct investment as a share of GDP (plus), net fuel exports as a share of GDP (plus), dummy for Middle East oil exporters (minus), and dummies for the 82 less developed countries and the three categories of FCP economies (uncertain). All of the continuous variables had the expected signs and were significant at the 90 percent level or higher. The model explained approximately two-thirds of the variation in the annual growth rate of per capita GDP of the 122 countries during the 15- and 20-year periods.

The economic freedom variable was included in two forms: (1) 1995 EFW summary rating and change in the EFW rating from 1995 to 2015 and (2) average EFW summary rating during the 1995–2015 period. The economic freedom variables were always positive and significant (in most cases, at the 99 percent confidence level). This indicates that economic freedom exerts a positive and highly significant impact on economic growth, even after accounting for the other factors included in the model. The dummy variable for the FCP group with a 2015 EFW rating above 7.5 was always significant, indicating that the growth rates of these seven countries were more rapid than the world's 21 high-income countries. The regression analysis provides additional evidence that (1) economic freedom exerts a strong impact on economic growth and (2) the FCP economies liberalizing the most grew more rapidly than those that liberalized by lesser amounts.

In addition to income, it is also important to analyze the factors underlying life satisfaction. Regression analysis was used to examine the determinants of life satisfaction, as measured by the World Values Survey (Institute for Comparative Survey Research 2017).

⁵The female population was chosen instead of total population because it more accurately reflects the latent composition of the population, which is sometimes contaminated by in-migration of workers, most of whom are male.

A set of personal attributes (such as employment, relative income, gender, and age) and country-specific measures including the summary EFW rating, per capita GDP, the Polity IV democracy score, and language fractionalization were incorporated as independent variables. The results indicate that economic freedom exerts a significant positive impact on life satisfaction both directly and indirectly (through per capita GDP). While the life satisfaction of persons living in FCP countries was well below that of similar individuals in other countries during the 1990s, the gap has declined, and by 2010–2014 it was virtually eliminated.⁶

Area Ratings and Identifying the Strengths and Weaknesses of the FCP Economies

In addition to the summary rating, the EFW data provide country ratings for five areas: (1) size of government, (2) legal structure and protection of property rights, (3) access to sound money, (4) international exchange, and (5) regulation of credit, labor, and business. The area ratings provide insights on both the strengths and weaknesses of economies. They also make it possible to track the source of changes in economic freedom of the FCP economies and compare their ratings with other European countries.

Table 4 provides the mean area ratings in each of the five areas for both the FCP economies and the 16 high-income European countries during 1995–2015. Looking at the mean ratings for Areas 1, 3, 4, and 5, we find that, in each of these areas, the mean rating of the FCP economies rose substantially during 1995–2015; their ratings also improved relative to the 16 high-income European countries. The high-income countries have low ratings in Area 1 (size of government). Thus, in this area, the mean rating for the FCP economies was higher than the mean for the high-income European countries. Moreover, the difference expanded during the two decades. In areas 3, 4, and 5 the mean ratings of the FCP countries were persistently lower than those of the European 16. However, the mean rating of the FCP countries rose steadily throughout 1995–2015 and the gap compared to the high-income European group narrowed. In Area 3

⁶The regression analysis of life satisfaction can be found in Table 16 of the full report, available at <https://coss.fsu.edu/hilton/sites/default/files/1b.%20FCP%20economies%20-%20Complete%20Study.pdf>.

(access to sound money) the rating improvement was huge and the narrowing of the gap dramatic. In 1995, the mean rating of the FCP countries was only 3.27 compared to 9.63, a gap of 6.63 units. By 2015, however, the mean Area 3 rating of the FCP countries had risen to 8.75 and the gap narrowed to only 0.76 units. While the gains were smaller for areas 4 (international exchange) and 5 (regulation), the pattern was the same: the mean rating of the FCP group rose substantially and the gap compared with the high-income European countries narrowed.

TABLE 4
MEAN AREA RATINGS FOR THE 25 FORMER CENTRALLY
PLANNED ECONOMIES AND THE 16 HIGH-INCOME
EUROPEAN COUNTRIES, 1995–2015

Area	Set of countries	1995	2000	2005	2010	2015
Area 1	25 FCP	4.46	5.48	6.32	6.17	6.26
	16 European	4.00	5.04	5.38	4.76	4.94
	Gap	-0.46	-0.43	-0.95	-1.41	-1.32
Area 2	25 FCP	5.68	5.81	5.45	5.60	5.48
	16 European	7.81	8.08	8.13	7.91	7.88
	Gap	2.13	2.27	2.68	2.31	2.40
Area 3	25 FCP	3.27	6.40	8.10	8.58	8.75
	16 European	9.63	9.52	9.48	9.34	9.51
	Gap	6.36	3.12	1.38	0.76	0.76
Area 4	25 FCP	7.36	7.55	7.25	7.33	7.80
	16 European	8.83	8.98	8.16	8.01	8.14
	Gap	1.47	1.43	0.91	0.68	0.33
Area 5	25 FCP	4.89	6.41	7.01	7.21	7.39
	16 European	6.70	7.58	7.80	7.63	7.98
	Gap	1.81	1.17	0.79	0.42	0.60

NOTE: The five areas are: (1) size of government; (2) legal structure and protection of property rights; (3) access to sound money, (4) international exchange; and (5) regulation of credit, labor, and business. See Table 3 for the list of the 16 high-income European countries.

SOURCE: Gwartney, Lawson, and Hall (2017).

Turning to Area 2 (legal structure and protection of property rights), we find that, in contrast with the other four areas, the mean rating of the FCP countries changed little in this area. The mean Area 2 rating of the FCP economies was 5.68 in 1995, 5.45 in 2005, and 5.48 in 2015. Further, the gap relative to the high-income European economies was 2.13 units in 1995, but it had expanded to 2.40 units in 2015.

Perhaps the patterns observed in Table 4 are unduly influenced by the FCP countries that have largely failed to move toward liberalization. In order to see if this is the case, the mean area ratings were also derived for only the 11 FCP countries that are now part of the EU.⁷ None of these countries was in the least-free group of the FCP countries. Thus, with only a few exceptions, these countries are the most economically liberal of the FCP economies.

When only these 11 FCP countries are considered, the pattern is the same as that of Table 4. The mean ratings of the 11 FCP economies increased substantially in Areas 1, 3, 4, and 5, and improved relative to the high-income European countries. But once again the situation for Area 2 was dramatically different. The mean Area 2 rating for the 11 FCP countries that are now EU members changed only slightly during the two decades. The mean Area 2 rating for this group rose from 5.97 in 1995 to 6.06 in 2005 and 6.09 in 2015. Moreover, the Area 2 mean rating of these countries was approximately two units less than the figure for the high-income European countries throughout the two decades.

Weakness in the legal structure area is a major problem for almost all of the FCP economies.⁸ Only one of the 25 FCP economies had a 2015 Area 2 rating above 7. Estonia's Area 2 rating in 2015 was 7.51, but the next highest Area 2 rating among the FCP group in

⁷These countries are Czech Republic, Estonia, Hungary, Latvia, Lithuania, Slovak Republic, Slovenia, Poland, Romania, Bulgaria, and Croatia.

⁸Given the historic background, the weakness of the legal systems of the FCP economies is understandable. Under communism, judges, lawyers, and other judicial officials were trained and rewarded for serving the interests of the government. Protection of the rights of individuals and private businesses and organizations was not important. Given this background, development of an independent legal system capable of checking the powers of the executive, protecting individuals and their property, and enforcing contracts even-handedly is a challenging task.

2015 was Georgia with a rating of 6.57. Only seven of the FCP economies (Georgia, the three Baltic countries, Czech Republic, Hungary, and Slovenia) had Area 2 ratings of more than 6 in 2015. Thus, 18 of the 25 FCP economies had Area 2 ratings of two or more units below the 16 high-income European countries. Moreover, there is evidence that the situation is worsening in several countries. For example, Poland's Area 2 rating in 2015 was 5.89, down from 6.21 in 2010. The Area 2 rating of the Slovak Republic was 5.78 in 2010 and 5.64 in 2015, down from 6.63 in 2005. Hungary's Area 2 rating fell from 6.66 in 2005 to 6.04 in 2015.

As we have shown, the FCP economies have grown rapidly and closed the income gap relative to the high-income countries of both Europe and the world. However, unless the deterioration in the legal structure of these countries is reversed and improved, it is unlikely these countries will continue to grow rapidly and close the income gap relative to high-income countries.

The legal system of a country is vitally important for sustained growth and achievement of a high per capita income. If investors—domestic as well as foreign—cannot count on protection of property rights and unbiased enforcement of contracts, they will be reluctant to undertake capital projects. In turn, weak investment will slow not only capital formation, but also entrepreneurial activities, dissemination of technology, and dynamic growth. There is already some evidence that this is happening in the FCP countries. Net FDI fell sharply during 2011–15. As Table 2 shows, the growth of per capita real GDP during the past five years has slowed. Perhaps these changes are caused by other factors, but they are precisely the outcomes one would expect from a poorly operating legal system.

Major Implications and Lessons for the Future

The fall of communism and the subsequent institutional change provide researchers with a natural experiment. The 25 countries involved in this analysis were diverse and they often chose different transitional paths. What are the most important lessons that can be learned from an examination of the changing institutional framework and accompanying performance of these economies? Four major factors stand out.

First, the experience of the FCP countries indicates that economic freedom enhances growth. The FCP economies that chose a course

more consistent with economic freedom grew more rapidly than those that were less free. Five of the seven most-free FCP economies achieved a robust annual growth rate of per capita GDP of more than 5 percent during the 20-year period from 1995 to 2015. The most-free group systematically achieved a higher mean growth rate than the middle and least-free groups. Further, consider the economic record of the 11 countries with the highest 2015 EFW summary ratings. These countries include late reformers such as Romania and Bulgaria, which did not begin the reform process until after 2000. During 2000–15, nine of these 11 countries achieved an annual growth rate of real per capita GDP of more than 4 percent. In contrast, only four of the 14 FCP countries with lower 2015 EFW summary ratings were able to achieve annual growth rates of more than 4 percent. Moreover, the high growth rates of two of the four (Kazakhstan and Azerbaijan) were elevated by the increasing and abnormally high oil prices during 2002–14. In addition, regression analysis indicates that, other things constant, economic freedom exerts a positive and statistically significant impact on growth of per capita GDP. This was true for countries throughout the world and for the most-free FCP economies.

Second, the FCP economies with more economic freedom experienced larger expansions in international trade and attracted more FDI than those with less economic freedom. The increases in the size of the trade sector of relatively free FCP economies such as Georgia, Lithuania, Latvia, Albania, Czech Republic, Bulgaria, Poland, and the Slovak Republic were truly remarkable. Even though FDI as a share of the economy receded during 2011–15, it was strong during the first decade of this century.

Third, during 1995–2015, the political institutions of most FCP economies moved toward the protection of civil liberties, democratic decisionmaking, and better control of corruption. The following nine countries had 2015 political institutions most consistent with civil liberties protection, political democracy, constraints on the executive, and absence of corruption: Estonia, Lithuania, Latvia, Czech Republic, Poland, Slovak Republic, Hungary, Croatia, and Slovenia. In contrast, the political institutions of Kazakhstan, Tajikistan, Russia, and Azerbaijan were least consistent with protection of civil liberties, democratic principles, and absence of corruption.

Fourth, the economic freedom area ratings of the FCP countries increased substantially in areas 1 (size of government), 3 (access to

sound money), 4 (international trade) and 5 (regulation of finance, labor, and business) during 1995–2015. The improvements in these four areas have narrowed the economic freedom gap of the FCP economies relative to the 16 high-income European countries. In these four areas, the economic freedom ratings of the FCP countries, particularly the eleven that are now members of the EU, are approximately the same as the ratings of the high-income European countries. However, there is a huge gap in the quality of the legal systems (Area 2) of the FCP countries compared to the high-income countries of Europe. Moreover, the FCP countries have failed to improve in this area. Unless the FCP countries improve their legal systems, their future growth is likely to slow and their gains relative to high-income countries come to a halt in the near future.

Looking back, the record of the FCP economies, particularly those that have moved toward economic liberalization, is one of growth, integration into the world economy, and increasingly democratic political decisionmaking. Looking forward, improvement in the legal systems of these countries is crucially important for the continuation of economic growth and rising living standards.

References

- Berggren, N. (2003) “The Benefits of Economic Freedom: A Survey.” *Independent Review* 8 (2): 193–211.
- Bjørnskov, C.; Dreher, A.; and Fischer, J. A. V. (2010) “Formal Institutions and Subjective Well-Being: Revisiting the Cross-Country Evidence.” *European Journal of Political Economy* 26 (4): 419–30.
- Dawson, J. W. (1998) “Institutions, Investment, and Growth: New Cross-Country and Panel Data Evidence.” *Economic Inquiry* 36 (4): 603–19.
- _____ (2003) “Causality in the Freedom-Growth Relationship.” *European Journal of Political Economy* 19 (3): 479–95.
- Faria, H. J., and Montesinos, H. M. (2009) “Does Economic Freedom Cause Prosperity? An IV Approach.” *Public Choice* 141 (1): 103–27.
- Feldmann, H. (2017) “Economic Freedom and Human Capital Investment.” *Journal of Institutional Economics* 13 (2): 421–45.
- Freedom House (2017) *Freedom in the World*, 2017 Report. New York: Freedom House.

- Gwartney, J.; Lawson, R.; and Hall, J. (2017) *Economic Freedom of the World: 2017 Annual Report*. Vancouver: Fraser Institute.
- Institute for Comparative Survey Research (2017) “World Values Survey.” Available at www.worldvaluessurvey.org.
- Marshall, M.; Gurr, T. R.; and Jaggers, K. (2016) “Political Regime Characteristics and Transitions, 1800–2016.” Polity™ IV Project. Available at www.systemicpeace.org.
- Nystrom, K. (2008) “The Institutions of Economic Freedom and Entrepreneurship: Evidence from Panel Data.” *Public Choice* 136 (3): 269–82.
- Pitlik, H., and Rode, M. (2016) “Free to Choose? Economic Freedom, Relative Income, and Life Control Perceptions.” *International Journal of Wellbeing* 6 (1): 81–100.
- Rode, M. (2013). “Do Good Institutions Make Citizens Happy, or Do Happy Citizens Build Better Institutions?” *Journal of Happiness Studies* 14 (5): 1479–1505.
- Transparency International (2016) “Corruption Perceptions Index.” Available at www.transparency.org/news/feature/corruption_perceptions_index_2016.
- World Bank (2017) “World Development Indicators.” Available at databank.worldbank.org/wdi.

BOOK REVIEWS

China's Great Migration: How the Poor Built a Prosperous Nation

Bradley M. Gardner

Oakland, Calif.: The Independent Institute, 2017, 219 pp.

China's rise in a generation from poverty and isolation to become a global economic power would have been impossible without the mass migration of people from the countryside to the city. In *China's Great Migration*, Bradley Gardner describes the role that this unprecedented movement of people has played in "the greatest development story in human history."

Market reforms in agriculture, trade, and industry all played major roles in China's emergence from the legacy of Chairman Mao, but the most underappreciated reform was the growing freedom of the Chinese people to move to cities to better their condition. Gardner tells this important story with sound analysis but also with firsthand accounts from his time on the ground in China.

The movement begins with the agricultural reforms of the late 1970s, which freed millions of Chinese workers from the shackles of collective farming. The rural laborers then migrated to the cities to work in the factories that had begun to produce labor-intensive goods for global markets. As a result, Gardner writes, "Between 1978 and 2012, the population of China's cities grew by half a billion people, swollen by more than 260 million economic migrants moving to urban centers to look for new opportunities."

From 1980 to 2010, the population of Beijing rose from 9 million to 21 million, Shanghai from 11 million to 20 million, and most

Cato Journal, Vol. 38, No. 1 (Winter 2018). Copyright © Cato Institute. All rights reserved.

incredibly Shenzhen, the area surrounding Hong Kong, from 300,000 to 10 million. New York and London had also seen their populations swell during the industrial revolution, but not by nearly so much nor during such a short span of time.

The most direct impact of this great migration of people was on the migrants themselves. By moving to the city, their output and incomes rose immediately by 500 percent. The factory work was hard and the conditions far below western standards, but “for China’s poor, the difference is life changing,” Gardner observes. “An extra 8 to 9 dollars a day is the difference between struggling to buy clothes and struggling to buy a phone; between an empty field and a flush toilet; between illiteracy and a technical education.”

For China as a nation, the great migration has been responsible for an estimated 20 percent to 33 percent of its economic growth since reforms began. That amounts to an additional \$1.1 trillion in economic output over a 20-year period. The number of people in China living in absolute poverty fell from 1981 to 2011 by more than 750 million, creating a middle class that has fueled China’s rise as the world’s largest consumer market for automobiles, luxury goods, and smartphones.

The great migration was not a deliberate part of the government’s development strategy. In fact, as Gardner documents, China’s communist authorities had traditionally tried to thwart the movement of its citizens. Beginning in 1951, the Chinese government enforced a household registration system known as *hukou*, which required residents to obtain the permission of police before they could move. Under Mao, *hukou* was “uniquely disastrous,” trapping peasants in the countryside just as the Great Leap Forward of 1958–62 was starving them of food.

As reforms took hold in the 1980s, China’s government generally looked the other way as tens of millions of technically illegal migrants moved to the cities. It was only in August 2003 that the *hukou* system was effectively abandoned when then Prime Minister Wen Jiabao abolished the accompanying law that allowed police to detain and repatriate migrant workers. “With one announcement,” Gardner writes, “the Chinese people were now free to move around their country without needing government permission.”

Remnants of the *hukou* system remain in place, but as a soft rather than a hard restraint. Many migrant workers still live a kind of second-class existence, free to work but not able to tap into

government services such as education for their children that are available to more long-time residents. Even so, the large majority are still much better off for moving.

Gardner's book is more than an academic work. A veteran of the Foreign Service and the Economist Intelligence Unit, he speaks Chinese and has lived and traveled extensively in China. He interviews migrants themselves and describes telling scenes such as the train stations crowded with Chinese workers headed home to the countryside during the Chinese New Year. This is a bottom-up story of the Great Migration.

Another strength of the book is how Gardner links the great migration to other reforms. For example, fiscal reforms in 1994 allowed the central government to collect more revenue but left local governments responsible for schools, local infrastructure, and social spending. To raise revenue, local governments engaged in land expropriation, undercompensating farmers for land on the fringes of urban areas and then reselling it at huge profit to developers to build industrial sites and high-rise apartments. While unjust, this system did open up land for development in a way that was accommodating to the incoming migrants.

For China, the great migration made possible its rise to middle-income status, but it has also covered up China's demographic decline and its inefficient use of capital. Because of its one-child policy, China's total working age population has already begun to decline. The internal movement of people peaked in 2007 and is expected to run its course by the mid-2020s, when there will be no more excess labor from the rural sector to fuel its industrial growth.

Thus time is running out on China's current model. "More people have migrated under the Great Migration than ever before in history but rural migrants are still a limited resource," Gardner writes. "China has yet to make the transition from an economy based on mobilizing labor to an economy based on capital efficiency. If it fails to make this transition in the next decade, it may run out of room to grow."

Gardner notes that even with all its reforms and progress, China's economy remains burdened by too much state intervention. To avoid a hard landing, "China needs to finish the reforms that it started more than thirty years ago. That means introducing full rural land rights, privatizing [state-owned enterprises], ending the *hukou* system, and most importantly making sure that local governments have

legitimate means to pay for legitimate costs. Absent these reforms, growth in China could quickly stagnate as debt and demographics take their revenge.”

For the rest of the world, China’s great migration reminds us once again that the freedom to pick up and move from one place to another is essential for human happiness and progress. This can be the freedom to move from a poor to a rich country, or from the countryside to the city within a country. In both cases, workers can experience a sharp increase in their productivity, benefiting themselves and their families while boosting total economic output.

“Giving people the opportunity to cross borders and pursue economy opportunities creates wealth for migrants, those around them, and the global economy. The largest barrier to migration is political, but political opinions can change rapidly,” Gardner concludes. “Fifty years ago, a common labor market in Europe would have been unthinkable; there is no reason that changes on a similar scale could not happen globally over the next fifty years.”

Daniel Griswold
The Mercatus Center
George Mason University

Taxing the Rich: A History of Fiscal Fairness in the United States and Europe

Kenneth Scheve and Daniel Stasavage

Princeton, N.J.: Princeton University Press, 2016, 259 pp.

Kenneth Scheve and Daniel Stasavage’s *Taxing the Rich: A History of Fiscal Fairness in the United States and Europe* is an attempt to investigate the politics of inequality through the lens of taxation. The authors place themselves in explicit dialogue with recent left-leaning scholarship on the economics of inequality, especially the work of Thomas Piketty and Emmanuel Saez. Piketty’s thesis in *Capital in the Twenty-First Century*, which appeared in 2013 and was an instant bestseller, is that capitalism stokes an ever-widening inequality of wealth and income due to the tendency for returns on capital to exceed returns to labor. He argued that only a scheme of highly progressive income taxation can halt or reverse this inherent source of inequality.

Scheve and Stasavage ask: What prompts democratic governments to heavily tax the rich? Their inquiry is relevant to both supporters and critics of Piketty. For supporters of income and wealth redistribution, the relevant issue is: What arguments or political conditions are needed to produce the kind of tax policies necessary to check inequality in a capitalist economy? For detractors, it is: What arguments or conditions might lead a democracy to undermine free markets, the incentives for productive economic activity, and individual rights for the spurious goal of reducing gaps in wealth and income? Scheve and Stasavage are especially interested in whether increases in inequality will prompt democracies to increase taxes on the rich. Like Piketty, they focus on a small slice of top earners by examining and emphasizing top marginal rates. In the contemporary United States, for example, about only 1 percent of households are in the top income bracket.

Taxing the Rich begins with a review of political thought about taxation in the first two chapters, then moves to a statistical analysis demonstrating the relation between high taxes and war mobilization in the third and fourth chapters. The fifth and sixth chapters look at the reasons behind the relationship between high taxes and war, and the final three chapters analyze the decline in tax rates from their midcentury peak and the implications for the future.

Scheve and Stasavage use positivist social science methods in their research. Positivism in the social sciences can be thought of as an attempt to apply the research methods of hard science to the study of social phenomena. As with the hard sciences, positivist social research involves empirically testing the causal relationship between a set of independent variables and a dependent variable. Such tests may yield generalizable social laws (nomothetic relationships) that hold true across a given set of societies, cultures, or time periods. The idea that democratically governed countries will not go to war with one another is, for example, a well-known product of positivist social science research.

Scheve and Stasavage's independent variables include democratization, inequality, state capacity, size of government, the nature and amount of government expenditure, and mass mobilization for war. The dependent variable in *Taxing the Rich* could be better defined, but in practice it consists of marginal tax rates that exceed 50 percent. The authors find mass war mobilization is the sole factor

driving marginal taxes above 50 percent and, by building a data set featuring modern, advanced democracies since 1800, it is suggested that their results should hold in similar kinds of advanced democracies in the future.

Scheve and Stasavage connect their statistical inquiry with a typology of theoretical principles used to justify different tax systems: (1) the ability-to-pay principle, (2) the equal treatment principle, and (3) the compensatory principle. “Ability to pay” is the idea that people’s tax burdens should reflect their capacity to pay, as measured, depending by the tax regime at issue, by income or wealth; “equal treatment” is the idea that all people should bear similar proportional tax burdens; and “compensatory” is the notion that taxes should compensate for the state’s role in favoring or harming certain groups of citizens relative to others, especially in the context of wartime drafts. Each principle reflects a different understanding of what constitutes fair treatment of citizens in a democracy.

In advanced democracies, they argue, all tax systems and arguments about taxes can be categorized according to these three principles. Not every system or argument reflects just one principle, but it is easy to separate out each constituent part from the whole and to determine the relative importance of each principle in motivating the structure of a given tax system. Their goal is to match the statistical results of their positivist inquiry with the three principles, which serve as reasons that can explain the correlations. The principles also provide context for thinking about generalizability, granting Scheve and Stasavage a basis for making a claim about the conditions likely to enable future high taxation of the rich.

After the chapter “treating citizens as equals,” Scheve and Stasavage review classic theoretical writing on taxation—including Arthur Pigou, Francis Edgeworth, John Stuart Mill, and Edwin Seligman—in order to present the ability to pay, equal treatment, and compensatory principles. The ability-to-pay principle is traced from the distinction between needs and wants in the 18th century Enlightenment to the marginal revolution in economics that occurred in late-19th century Austria. The term “compensatory theory” comes from the Progressive era American economist Edwin Seligman, who, in his 1911 book *The Income Tax: A Study of the History, Theory, and Practice of Income Taxation at Home and Abroad*, used it to refer to the idea that progressive taxation is justified in instances where high income or wealth primarily reflects

luck. Scheve and Stasavage home in on a more narrow meaning, referring only to instances where taxation is used to remedy privileges given by the state, their main example being exemption from wartime drafts. They spend much less time (only a few lines) on equal treatment, the idea that public-finance scholars more commonly refer to as horizontal equity.

After laying out these intellectual foundations, Scheve and Stasavage move to data analysis. They succeeded in compiling a comprehensive data set of the annual top marginal income tax rates for 20 democracies from 1800 to the present. In the third and fourth chapters they test various hypotheses for possible causal factors that might determine upward movements in the top rate. Scheve and Stasavage demonstrate that mass war mobilization is significantly correlated with increases in top marginal income tax rates, while other possible factors, like democratization, leftist political influence, and inequality, are only weakly related.

The authors' next goal, pursued in the fifth and sixth chapters, is to determine why taxes on the rich increased in such close tandem with mass war mobilization. This is a much harder task than establishing the correlations and, unfortunately, Scheve and Stasavage's methods are less praiseworthy and their arguments less convincing. The authors argue that mass mobilization led policymakers, thought leaders, and the public to agree on compensatory tax principles. However, only two pieces of evidence are offered to defend this claim. First, the authors point to public opinion polls in the United States during World War II that show broad support across all socioeconomic groups for significant increases in the effective tax rate paid by the rich. Second, they identify similarly wide support in the British parliament and among U.S. intellectuals during World War I for the compensatory principle. Scheve and Stasavage do not devote much space to discussing these pieces of evidence. Rather, they present other hypotheses that might explain the correlation between mass mobilization and high taxes on the rich and attempt to debunk them.

Therein lies the weakness in positivist social science research: while such methods can be very helpful in establishing general links, pure statistical analysis is less useful for the purpose of offering a comprehensive explanation of a complex social phenomena. The problems with the monocausal premise are on display in Scheve and Stasavage's attempt to negate the importance of nationalism and

fiscal necessity as possible causes of the wartime tax structure. In both instances the authors claim these possible determinants lack any salience only because they do not work as catch-all explanations.

Scheve and Stasavage correctly argue that nationalism alone would not account for high taxes on the rich. The authors stop there, however, dispensing with the discussion of nationalism or other cultural issues. Essentially, Scheve and Stasavage dismiss the importance of a possible causal factor because it is not sufficient on its own to explain high tax rates. Regarding fiscal necessity, they reject that notion because wartime tax structures were never optimal for the purposes of revenue collection. If revenue was the abiding concern behind wartime tax policy, rates would have been higher across the board. But as with nationalism, this conclusion is quite simplistic. At least in the literature on American tax history, revenue often is cited as the driving concern of fiscal policy during both world wars. Surely the need for revenue and concerns about inflation contributed to the acceptance of the wartime tax structure as “fair.”

As a result of the deficiencies in their approach, Scheve and Stasavage fail to make a convincing case that high taxes on the rich during mass mobilization mainly reflected the compensatory principle of using taxes to remedy the relative ability of the rich to avoid the draft. That Scheve and Stasavage are likely missing important parts of the story can be seen from a closer examination of the relation between the state and society in America during World War II. That topic was the subject of an excellent book by James Sparrow, *Warfare State: World War II Americans and the Age of Big Government* (reviewed in the Spring 2012 *Cato Journal*). Sparrow convincingly shows that mass participation, rather than high taxes on the rich, was the defining characteristic of the World War II fiscal state. Sparrow explains that three primary concerns drove tax policy: increasing revenue; checking the wartime inflationary gap by soaking up discretionary income; and furthering civilians’ psychological attachment to the war effort. Americans’ fiscal relation to the state changed dramatically as a result. In 1939, only 7 percent of the labor force paid federal income taxes; in 1944, 66 percent did. While there are no hard statistics on the number of bond owners, Sparrow claims that bond ownership was “near universal.” Tax payments and bond ownership served to create a sense of fiscal citizenship, or at least a widespread feeling that paying taxes made one a member of the national polity.

If World War II tax policy reflected mass feelings of involvement and sacrifice, why were taxes on the rich so high? The Revenue Act of 1942, for example, set a rate of 88 percent on marginal income above \$200,000. Sparrow notes that most new filers' tax payments were essentially voluntary because of limited federal administrative capacity and that tax burdens were steep across the income distribution, including for low-income payers. Perhaps very high taxes on the rich were needed to maintain the sense of egalitarianism behind the mass acceptance of the tax system. Or perhaps high-income taxes reflected a society-wide phenomenon where citizens identified with the plight of the soldier by limiting consumption.

Scheve and Stasavage's final three chapters analyze the path of tax rates in advanced democracies after World War II and offer predictions about the possibility for increases in the top graduated rate in the future. They argue that top tax rates fell as memories of the wartime social dynamic gradually faded. They suggest that taxes on the rich in advanced democracies will not significantly increase in the future unless another instance occurs where government openly and systematically privileges the rich in a matter central to public life. As warfare no longer requires mass armies and a draft, it is difficult to imagine such a scenario occurring. Scheve and Stasavage's arguments that a social contract akin to what took root in the world wars is unlikely to emerge in the future are logical and convincing. However, they are dismissive of economic efficiency arguments in explaining the decline in taxes after World War II, again showing that their emphasis on a single, driving cause of top tax rates compromises the validity of much of their thinking.

In sum, while Scheve and Stasavage are persuasive that mass mobilization correlated more closely with high taxes on the rich than any other discernable factor, they do not provide enough evidence to show how such high taxes resulted mainly from the compensatory principles of using the tax code to force the rich to sacrifice. Therefore, their argument that heavy taxation of the rich in democracies is determined by the strength of compensatory arguments fails. For better or worse, observers of public affairs should not expect that Piketty's steeply progressive tax system will be implemented only if the social contract systematically favors the wealthy.

Finally, the inquiry into the causes of high wartime taxes should not have tried to find a statistical smoking gun. Instead, Scheve and

Stasavage should have attempted to produce more insights backed by heavy reasoning. Such insights could follow from a rigorous description of the politics of taxation coupled with a careful application of theory. With the attention received by Piketty, the ascendancy of the Bernie Sanders left, and Steve Bannon's recent proposal to increase the top marginal rate to above 40 percent, smaller insights grounded deeply in qualitative evidence would go much further in helping us make sense of the current tax politics paradigm than Scheve and Stasavage's bold, but ultimately hollow, analysis.

Ari Blask
Columbia Law School

Churchill and Orwell: The Fight for Freedom

Thomas E. Ricks

New York: Penguin Press, 2017, 349 pp.

At first glance, a dual biography of Winston Churchill and Eric Blair—who went by the pen name George Orwell—sounds like a project that shouldn't be pursued. After all, Churchill was an upper-class imperialist politician who disdained socialism. Orwell, a committed socialist, was a struggling novelist and journalist for most of his life. The two men never met each other and moved in very different circles.

And yet both men were among the few to correctly foretell the dangers of the two greatest threats to human freedom in the 20th century: nazism and communism. This shared prescience and the two men's steadfast and oftentimes lonely opposition to these two tyrannies are the focus of Thomas E. Ricks's latest book, *Churchill and Orwell: The Fight for Freedom*.

Ricks's book opens by noting another similarity between the two men: they were both almost killed before achieving notoriety. During the Spanish Civil War, a Nationalist sniper shot Orwell in the neck while he was fighting with the anti-Stalin POUM (Workers' Party of Marxist Unification). Miraculously, the bullet missed his spinal cord, arteries, and windpipe. Orwell also survived the Soviet-backed repression of the POUM.

A New York City cab driver struck Churchill at around 30 miles an hour in 1931. At the time, Churchill was enduring what he called his "wilderness years" of political isolation and was in America

hoping to pick up some speaking fees following the financial crisis of 1929. The cab dragged Churchill down the street, cracking ribs and cutting his scalp. And, during the Second Boer War, a young Churchill survived a prisoner of war camp (that he later escaped) and was at the Battle of Omdurman in Sudan.

As Ricks notes, had Churchill and Orwell been killed in Spain, South Africa, Sudan, or New York they would be remembered by only a few historians specializing in minor British politicians and literary figures. Fortunately, Churchill and Orwell survived their brushes with death, going on to become two of the last century's most influential figures.

Achieving such influence was not, however, without significant obstacles. Churchill had to contend with broad support for appeasing Hitler. The First World War, which led to the deaths of 886,000 British military personnel, had only concluded roughly two decades before Germany invaded Poland on September 1, 1939. Many in Britain were hardly keen for another war.

Appeasement was the goal of Prime Minister Neville Chamberlain, whom Churchill succeeded in May 1940. But Chamberlain was hardly alone. As Ricks writes, "Dallying with fascists was not simply a pursuit of the young and the foolish." Some members of the Houses of Lords and Commons, the *Times* of London, and King Edward VIII all supported appeasement. Hard as it is for many to believe today, Churchill's warnings of Germany's rearmament fell on many deaf ears.

Orwell's voice was also a whisper in a whirlwind at times. While today considered almost a prophet of totalitarianism, Orwell faced criticism for his unrestrained opposition to the Soviet Union. *Animal Farm*, his classic anti-Stalin fable, struggled to find a publisher. And, although Soviet operatives tried to prevent the book from being published, it was eventually published in Britain in 1945, with Harcourt Brace buying the U.S. rights to *Animal Farm* for 250 pounds in December 1945 and publishing in 1946. Since then, with millions of copies have been sold around the world.

Ricks respectably manages to convey an admiration for Churchill and Orwell without ignoring their faults. He discusses Orwell's early anti-Semitism, his naïve faith in centralized economic planning, and the fact that he compiled a list of suspected communists for the British government. And, although Churchill won a Nobel Prize in literature, Ricks directly addresses the numerous factual errors in his

epic six-volume memoir *The Second World War* and his four-volume history *A History of the English Speaking Peoples*. He also mentions that Churchill's second stint as prime minister from 1951 to 1955 was hardly a glowing success. As Churchill biographer Roy Jenkins noted, "It is impossible to re-read the details of Churchill's life as Prime Minister of this second government without feeling that he was gloriously unfit for office."

Unfit for peacetime leadership Churchill may have been, but the greatest conflict in human history suited him very well. He understood the stakes, declaring in a September 1939 speech in the House of Commons, "This is no war of domination or imperial aggrandizement or material gain; no war to shut any country out of its sunlight and means of progress. It is a war, viewed in its inherent quality, to establish, on impregnable rocks, the rights of the individual, and it is a war to establish and revive the stature of man."

Later in the same speech Churchill noted that the House of Commons had been voting on wartime bills stripping the British people of some of their liberties, saying, "We look forward to the day . . . when our liberties and rights will be restored to us." But rights were not restored after the Second World War. Churchill's socialist successor, Clement Attlee, continued the national ID scheme implemented by the National Registration Act of 1939. Parliament repealed the act seven years after the war ended. The last person to be prosecuted for not producing his ID card was the liberal hero Clarence Harry Willcock, who told a London policeman, "I am a liberal, and I am against this sort of thing."

Churchill understood the risks of the war as well as the stakes, noting that his own commitment to fighting Hitler could lead to his death. As he told colleagues in a cabinet meeting, "If this long island story of ours is to end at last, let it end only when each of us lies choking in his own blood on the ground."

Orwell also felt that his firm stance against tyranny endangered his life. Following the publication of *Animal Farm* he bought a pistol to protect himself from communists. As Orwell knew from firsthand experience in Spain, Soviet agents had no qualms when it came to assassinating dissidents.

Churchill and Orwell can be praised for their foresight and bravery, but is either man worth paying attention to now? Libertarians in particular may balk at the prospect of taking lessons from a socialist and an imperialist.

And yet when anyone on the political spectrum today discusses government snooping, they speak in a language forged in Orwell's novels. "Big Brother," "thought police," "doublethink," the "memory hole," and "Room 101" are all plucked from *1984*. *Animal Farm* includes its own phrases that have found their way into modern speech, perhaps most notably, "All animals are equal, but some animals are more equal than others." Indeed, Orwell's own name has become an adjective, with modern surveillance programs being described as "Orwellian."

Orwell also emphasized how governments engage in linguistic perversion in order to confuse meaning and help achieve their ends. In *1984*, the English Socialist Party proclaims, "War is Peace, Freedom is Slavery, Ignorance is Strength." This is only one of the examples in the novel of the government asserting contradictions. The Ministry of Truth, where the protagonist Winston works, is the government's propaganda mill, erasing history. The ministry that oversees torture is the Ministry of Love. No one reading Orwell's work today can avoid thinking of contemporary examples. The United States doesn't bomb enemy fighters and torture people; it launches "kinetic military actions" against "enemy combatants" and uses "enhanced interrogation techniques."

Churchill himself is better known for phrases and quips rather than pieces of vocabulary, and Ricks separates the mythology of Churchill's wit from reality. Churchill did say, "I've taken a lot more out of alcohol than it's ever taken out of me," but he didn't tell Lady Astor that he would drink poisoned coffee if he were married to her.

Alcohol-related banter aside, Churchill is renowned for his speeches that still inspire today. "We will never surrender," "Never in the field of human conflict was so much owed by so many to so few," and "I have nothing to offer but blood, toil, tears, and sweat" are but a few examples.

Ricks dedicates a section of one chapter to Churchill's occasional open displays of tears during the war, something that runs contrary to the British people's aversion to public expressions of emotion. Those tears sometimes appeared when Churchill delivered speeches he had written himself, a testament to his writing talents (or his ego, perhaps). Like Orwell, Churchill's mastery of the English language helped his name become an adjective: "Churchillian."

As things stand, our governments are not Orwellian (the Patriot Act and GCHQ notwithstanding) and our political rhetoric is hardly

Churchillian. Those with a propensity to write “We live in an Orwellian dystopia” would not be able to do so in truly totalitarian states. While government surveillance is a serious concern, we’re not anywhere close to living in Orwell’s Oceania.

A glance at the current crop of political leaders uncovers no one with Churchill’s rhetorical skills. That is a shame, especially given that Ricks’s book leaves one feeling that a stubborn political orator with a penchant for drinking, the written word, and human freedom would be very welcome right now. So too would a journalist with Orwell’s writing style, combat experience, skepticism of authority, and imagination.

Aside from exhibiting his subjects’ talents and flaws, all of which have been discussed for decades, Ricks provides some valuable insights to anyone concerned with the state of liberalism. When it comes to the defense of freedom, allies can come from any wing of the political spectrum, but mounting a defense of liberalism still can be a lonely—though ultimately worthwhile—endeavor.

Matthew Feeney
Cato Institute

The Limits of the Market: The Pendulum between Government and Market

Paul De Grauwe

New York: Oxford University Press, 2017, 165 pp.

In the preface to *The Limits of the Market*, Paul De Grauwe, an economics professor at the London School of Economics, begins with two basic premises: first, that a centrally planned economy does not work; second, that pure market systems do not exist anywhere. According to De Grauwe, “The only relevant question, then, is how precisely that mixture should look”—that is, the correct mix of market freedom and public policy oversight.

He writes of a “Great Economic Pendulum”—that swings between greater and lesser government intervention—“cyclical movements” in economic history. To illustrate this, he describes a general decline in government involvement in the world economy during the first part of the 20th century. Then the Great Depression ushered in greater government involvement that continued through the middle decades of the century. That gave way to renewed

market liberalization in the 1980s, which ended with the Great Recession. With this brief overview of modern economic history as his departure point, he begins to explain what he sees as the limits of capitalism.

“The limits of a market system relate to the fact that the connection between individual and collective rationality can be severed,” De Grauwe writes. He identifies two types of situations in which large numbers of people do not believe that their own interests are being met satisfactorily by markets. The first type, which he labels “external limits,” and others commonly call “externalities,” involves individual decisions that have a positive or negative effect on others. The second category, “internal limits,” involves the conflict between individuals’ “System I” decisionmaking—relying on intuition and emotion—and “System II” decisionmaking—relying on reason. These two systems are connected and require a satisfying balance between emotional and rational demands in order for the market system to work effectively. Often, however, this balance of individual and collective rationality is not maintained.

DeGrauwe argues that individuals do not take into account the external effects of their individual decisions. He offers three examples of this: climate change and global warming (“carbon dioxide as a harmful gas generating external costs across geographic borders”); financial markets (“a banking crisis where the system reaches its limits before collapsing and requiring a government bail-out”); and public goods and the free-rider problem (“the market system has no mechanism for creating public goods”).

He discusses how the expansion of the market system brings it to its “internal limits.” He identifies three mechanisms that lead to the discrepancy between individual and collective well-being.

The first mechanism has to do with markets and distribution, whereby the price system leads to an inequality of income and wealth, a sense of injustice, and a desire for individual charity on the part of those able to purchase a good (“sharing with those not able to pay for a product”), as well as a possible revolt by those individuals in the lowest income groups (“protests and violent revolutions”).

The second mechanism has to do with intrinsic and extrinsic motivation. Intrinsic motivation means “that people are motivated to put their effort into their work or other activities because they are fulfilled by the work or activity itself.” Extrinsic motivation means “that people make an effort for the sake of the financial reward for a

particular activity.” Reconciling these two diametrically opposed motivations is inherently challenging, as managers find it easier to focus on extrinsic motivations involving financial bonuses (which are easily eliminated when managers are confronted by company financial duress).

The third mechanism has to do with competition and cooperation in the market environment. Though both can encourage good market behavior, competition and cooperation can conflict with each other on both emotional and rational levels, resulting in an internal tension between individual and collective well-being.

De Grauwe next addresses what he refers to as the “utopia of market self-regulation.” While so-called market fundamentalists argue that problems concerning the environment and public goods are grounded in a lack of property rights and solutions lie in somehow creating those rights, De Grauwe questions how such rights are possible. For instance, how can government establish ownership rights when people all over the world are negatively affected by air pollution? How does one identify billions of harmed people who will seek legal compensation from billions of other people who are responsible for their injuries, and in what court of law (De Grauwe calls this a “complicated information problem”)? But, as he notes, even if one were able to solve this information problem—that is, through the highly unlikely scenario of negotiating voluntary agreements with billions of people—the legislative branch will have to assign property rights in law (or negotiate those voluntary agreements), and the courts and police will have to enforce those rights—thus inevitably involving government institutions.

Given this need for government to ultimately save the market from its own destruction, De Grauwe discusses three domains where government policy and normative theory have a role in managing the market: externalities (by levying taxes on those engaging in undesirable behavior), supplying public goods (and eliminating the free-rider problem in market responses), and redistribution (of income and wealth on the highest incomes, thus providing the socially and politically stabilizing effects of greater equality). However, he recognizes that, while the government may know what it must do, that does not always mean that it will implement such policies. Why? Because government is also subject to limitations.

According to De Grauwe, politicians are prevented from following the normative prescriptions of economic theory. The external

limits of government, he writes, involve externalities related to the collective interest, say for environmental protection. This is a similar discrepancy that exists with the external limits of markets, but, unlike the market where individuals resist the efforts of governments to harm private interests in favor of the collective, the discrepancy works in the opposite direction: individuals make decisions that harm the collective good. In this case, the challenge for government (especially in democratic societies) is to close the difference between individual and collective rationality, thus gaining broader support from their citizens for the government's decisions.

As De Grauwe points out, the free market appeals to the individual's rational, calculating side at the expense of his or her emotional side. Thus, government appeals to the emotional side—or the dissatisfaction gap left by the free market—and is the mechanism through which emotions are expressed. To this end, government focuses its efforts on distribution problems, motivated by a universal sense of fairness. Emotional appeals to “fairness” win out over rational appeals to profit and efficiency.

Government, however, faces two limits. The first—redistribution at the expense of efficiency—holds that neither excessive equality nor excessive inequality is good for efficiency. Most capitalist regimes tend to reflect this principle. The second—a balance between emotion and reason—recognizes that the more governments work on redistribution the more they tax individuals and limit individual freedom. An example of this is embodied in a public-pension system, where, if government does not keep moral hazard under control, the citizen's sense of fairness is overwhelmed by too many “free-riders” exploiting the system.

De Grauwe notes that discussions on the role of markets and government have been ongoing for centuries “between people who believed that the market system was the foundation of economic organization (and creates the economic value that makes it possible to maintain the public sector) or that the task was reserved for government (as it establishes and enforces property rights).” That traditional view is espoused by both market fundamentalists and government fundamentalists, and is often portrayed as a pyramid hierarchy, with either the market or government at the top. So who is in charge?

De Grauwe argues that such hierarchical thinking is wrong, as both the private and public sectors are necessary and it is equally important for both sectors to operate effectively. To that end he

offers a nonhierarchical, elliptical portrayal of the relationship of market and government. As he explains:

If the government does not succeed in making those who generate external costs from market activities pay for them, then at some point the market system will come to an end. In the same way we can conclude that a government which does not succeed in providing essential public goods such as infrastructure, law and order, public safety, and education will lead the way to the downfall of the market system

The only relevant question worth asking, he writes, is how the division of labor between markets and government can best be organized.

De Grauwe next reviews theories of the linear historical demise of capitalism proposed by Karl Marx, Rosa Luxemburg, Vladimir Lenin, Joseph Schumpeter, and Karl Polanyi. In response to their critiques, he explains that the empirical data of the last 200 years have shown a cyclical—not linear—movement. The market system hits its limit, wanes temporarily, and is replaced by a state system that is full of internal contradictions that will lead to its demise, eventually to be replaced by a resurrection of the market system.

De Grauwe also views the creation of the eurozone (and by extension, the euro) as a threat to national governments. He explains how a eurozone government issues debt in a currency over which it has no control. Thus, the national government cannot offer any guarantee to bondholders that the euros will be available to pay them on the maturity date. Thus, without a national guarantee, the implication is that the financial markets can force national governments into insolvency against their will, which will lead to a dangerous supremacy of the markets. De Grauwe believes that the market has forced many eurozone countries (such as Greece, Portugal, and Spain) to adopt excessive austerity budgets that have led to increasing unemployment and the collapse of parts of the public pension system.

To better understand the limits of capitalism, he refers to the results of Thomas Piketty's *Capital in the Twenty-First Century*. Piketty begins his analysis by evaluating the long-term development of the quantity of (net) capital in the economies of the western world, focusing his historical analysis on France and Great Britain (although referencing the United States). He finds that, in the very long term,

capital income is higher than the growth of the economy, thus the return on capital is greater than the growth of gross domestic product. De Grauwe believes the implications of Piketty's research are reflected by the increasingly large proportion of top incomes in capitalist societies that go to the *rentiers*—the investors in capital. To save both democracy and capital, Piketty's solution is to offset the great inequality in *rentier* assets by instituting a progressive wealth tax, thereby saving capitalism from the capitalists.

In his concluding chapter, De Grauwe acknowledges there will always be pendulum swings between market and government. In the 21st century, he views the natural environment as an external limit toward which the market system is rapidly headed, and the tendency for greater income inequality as an internal limit of the market system. A reformist scenario, however, has governments raising taxes on high incomes and wealth, thus making inequality less extreme and more acceptable to society in the 21st century. By instituting this policy, governments would allow the market system to softly come up against its internal limits. Likewise, citizenry would apply democratic pressure on governments to combat pollution (such as carbon dioxide) by using taxes and regulations, and to provide subsidies for alternative energy sources. Therefore, this policy response halts environmental degradation and allows the market to more gently approach its external limits.

De Grauwe's basic framework offers a plausible explanation for how the pendulum swings between a mix of market freedom and government oversight and intervention. Modern economic history, with the cyclic swings in western democracies embracing capitalism bears this out. What that "mix" should be, however, is less convincing. He admits that he has evolved into less of a supporter of market forces, which is reflected in the title of his book. His argument is based on his belief that "neither market nor government is more important," yet he emphasizes examples of market failure far more than examples of government failure, though the latter includes such drastic examples as failed communist regimes or the potential for eurozone failure to back sovereign-issued debt. This book would have been improved if De Grauwe had dealt in finer detail with the true limits of government, for example by including the theory on nonmarket failure in democratic and capitalist societies as developed by Ronald Coase and Charles J. Wolfe, among others.

For example, De Grauwe uses the economic recession of 2008, specifically the financial services sector, as an example of market failure. The Financial Crisis Inquiry Commission, a U.S. commission charged with investigating the causes of the financial crisis, reported in 2011 that, while the modern financial system was evolving, government regulation did not keep pace with the rising risks in the U.S. economy. As the report states, “We had a 21st century financial system with 19th century safeguards.” That was not a market failure, but a government one—specifically, an egregious example of how the federal government performed inadequately and did not sufficiently intervene. This form of nonmarket failure is characterized as regulatory arbitrage, where financial institutions take advantage of the differences between economic risk and the existing regulatory position.

On the market side, De Grauwe discusses market failures in self-regulation but does little to explore rich examples of successful private ordering. Recent work by economist Edward Stringham (*Private Governance: Creating Order in Economic and Social Life*) offers a rational approach to addressing when government or markets are warranted to solve societal problems. Private governance is practiced by clubs and voluntary associations in the marketplace, through electronic payment processing, for example. Other examples include best industry practices, as well as codified industry standards, which are developed in the United States by associations without direct government involvement. (Such standards are often later incorporated by reference in administrative rules.) While often requiring a certain level of government intervention, for example, to adjudicate and enforce contracts, such private ordering is an example of how markets can operate creatively and effectively with limited government intervention and oversight.

In conclusion, De Grauwe offers the reader a thought-provoking book and a basic theoretical and empirical departure point for further work on an important topic relevant to the future of democratic capitalism. Public policy issues of income inequality (especially a stagnant or declining economic middle class), crushing national debt, public pension and health insurance funding shortfalls, and the ever intrusive and coercive regulatory state (and loss of liberty) are coming due. The important question still remains: What is the right mix of market and state?

Thomas A. Hemphill
University of Michigan–Flint

Oppose Any Foe: The Rise of America's Special Operations Forces

Mark Moyar

New York: Basic Books, 2017, 402 pp.

In recent years, special operations forces (SOFs) have assumed a prominent role in the ongoing U.S. global war on terrorism (though that term has fallen out of fashion). In the campaign against the Islamic State of Iraq and the Levant (ISIL), SOFs are fulfilling a number of significant responsibilities: launching periodic raids against high-value terrorist targets; providing tactical advice to partner forces from advanced positions; and coordinating fire support for lightly armed irregular forces on the front lines. Moreover, those activities extend beyond Iraq and Syria. A small SOF contingent is on the ground in Libya. The Pentagon has deployed an SOF task force to the Horn of Africa and Yemen, and special operators make up a significant portion of U.S. troops remaining in Afghanistan. Without a doubt, SOFs are enjoying something of a heyday within the U.S. military.

In *Oppose Any Foe: The Rise of America's Special Operations Forces*, Mark Moyar clearly demonstrates that the current stature of SOFs within the U.S. military is rather exceptional. He provides a comprehensive history of the institutional evolution of U.S. SOFs and their operational contributions from the Second World War to the present. In doing so, he paints a picture of an entity in search of a role. Largely unappreciated by the conventional military leadership through most of their history, SOFs have frequently been employed in situations for which they were ill suited and poorly equipped.

For readers who enjoy detailed narratives of military operations, *Oppose Any Foe* will be quite entertaining—although the inherent drama of those operations is frequently compromised by Moyar's penchant for bizarre similes that distract more than they illuminate. Moreover, some readers (myself included) would prefer a more detailed discussion of the politics driving the employment and institutional development of SOFs. One of the most interesting aspects of the history of SOFs is the “intense rivalry between special operations forces and regular forces,” which Moyar highlights as an enduring challenge. Unfortunately, skirmishes within the Pentagon receive much less attention than those on the actual battlefield.

Although *Oppose Any Foe* is a history, the book is clearly oriented toward the future, highlighting a number of important questions that have yet to be resolved. First and foremost, Moyar questions the extent to which SOFs should be expected and tasked to address the myriad security challenges currently facing the international community. In the wake of the massive interventions in Afghanistan and Iraq, numerous experts (both inside and outside government) have suggested that SOFs provide a means for effectively addressing a range of security challenges with a relatively “light footprint.” Particularly as the U.S. campaign to roll back the Islamic State has gradually succeeded, more and more observers have suggested that SOFs should be used in a similar fashion to combat future challenges “by, with, and through” local forces.

Yet Moyar contends that Obama “administration strategists had not given adequate consideration to the strengths and limitations of SOFs before hoisting them to the apex of the world’s most powerful military.” On a basic level, he highlights the necessity of identifying a sustainable level of activity for relatively small special operations forces. There is a broad consensus that the number and frequency of operations imposed on SOFs in recent years has strained readiness and morale within the units. As Moyar points out, since SOFs have already increased from 38,000 to 70,000 since 9/11, it is questionable whether they can either expand further or maintain the current operations tempo without degrading the elite capabilities that distinguish them from conventional forces.

Furthermore, the strategic utility of SOFs remains an open question. One theme that recurs throughout *Oppose Any Foe* is that SOFs have frequently achieved tactical successes but struggled to exert strong influence over strategic outcomes. For example, on first impression, the ongoing anti-Islamic State campaign would seem to represent an exception. SOFs have played a crucial role in dislodging ISIL from most of the territory it had seized in Iraq and Syria. On a broader political-military level, however, it remains to be seen whether the campaign will foster enduring stability and security—the absence of which invited the Islamic State onslaught in the first place. In other words, even when SOFs can win a war, it remains to be seen whether they can secure the peace.

An even more important question, however, is whether the United States’ current reliance on SOFs ultimately serves the national interest. It seems that one of the primary reasons the White

House has employed SOFs on such a significant scale in recent years is that they provide a convenient means for conducting military operations in numerous countries largely from the shadows. Deploying relatively small contingents of SOFs enables the White House to conduct a perpetual global war on terrorism without engendering much attention (or pushback) from the American public.

There are two risks to such an approach, however. The obsession with low visibility introduces a risk that SOFs will be employed in circumstances for which they are ill-suited and poorly equipped. Even more important, an overreliance on SOFs has the potential to subvert the American political process. The American public has the capacity to act as a powerful brake on ill-conceived military adventures. If a president can sustain a military operation only by hiding the United States' role from its citizens, one must question the ultimate wisdom of that enterprise. As Moyar suggests, "Presidents, being highly political animals, will continue to face temptations to use special operations forces to serve political agendas. For the good of the republic and the special operations forces, they would be well advised to resist those temptations."

Brad Stapleton
Alexandria, Va.

Against the Grain: A Deep History of the Earliest States

James C. Scott

New Haven: Yale University Press, 2017, 312 pp.

The subsistence farmer of cereal grains has been the most common human type in recorded history. Indeed, the 21st century may be the first recorded century during which subsistence farmers did not predominate. Or so we may hope.

But how did we become subsistence farmers? Recent archaeology suggests that the process was anything but easy and that states played a central role in the necessary subjugation of unruly waters, lands, plants, animals—and people. And a process of subjugation it most certainly was.

Early in recorded history—and in the tantalizing era that came just before it—something momentous took place, something that you probably never learned about in school: the state *invented* and *imposed* the subsistence farming of cereal grains. In the process the

state destroyed a way of life that is barely legible to us anymore, all but lost in the mists of time. Yet that way of life was in many respects superior to what followed. The peoples who were subjected to state power and subsistence fought against these impositions valiantly—and they lost. From that point on, the victorious state took care to keep almost everyone at the level of subsistence, or just below it, by appropriating nearly all the surplus product for itself. Although few of us have even realized it, the farming of cereal grains has been part of a nightmare from which we have all been trying to awake.

The story I've just told is highly unconventional, and yet it probably resonates with a certain strain of libertarian, even as it infuriates many others. It is also a very new story, which is why you didn't learn about it in school.

One place you will learn about it is in James C. Scott's latest book, *Against the Grain: A Deep History of the Earliest States*. It tells the story of a complex pattern of grain agriculture, early states, forced labor, and the extraction of surplus, and how all of these things were connected in ways that researchers previously never suspected.

In the conventional story, the one just lately overthrown by new discoveries in the Middle East and elsewhere, the first states had been forces for good. They rode in benignly on a crest of surplus grain and newfound human vigor. States then sponsored benevolent public works, like irrigation projects and cultic sites that celebrated the wonders of grain agriculture—to which humanity had already gravitated on its own. Cities became possible through surplus, and city life was also the creation of the state. With the advent of the state, living conditions improved rather than getting worse.

None of this, though, is altogether sympathetic to libertarians, who have always been critics of the state. And none of it is thought to be well supported by the evidence anymore. Earlier research into the first states, Scott argues, had been too prone to buy those states' own propaganda, which we must stop doing.

Armed with the latest in satellite and computer-aided archaeology, specialists now declare that the first farmers weren't always at subsistence. Before recorded history, and before the state, there had been not subsistence farming but flexibility and abundance. For many thousands of years—for longer than recorded history itself—humankind lived on a healthy mix of foraged plants, hunted animals, and low-intensity agricultural crops. Their labor was easy and short.

They were relatively healthy. They probably enjoyed more leisure time than we do today.

Nor was subsistence farming a natural development. Farmers were reduced to subsistence, and were kept there for centuries, through institutionalized theft, which is to say they were reduced to subsistence through taxation. Many populations appear to have resisted sedentary grain agriculture, resisted the state that imposed it, and resisted being confined to life in or near the unhealthy and disaster-prone new sites called cities.

Objections are easy to make. It's not hard at all to find similarities here to the thought of Jean-Jacques Rousseau, in that the French *philosophe* likewise believed that grain agriculture and civilization had been curses upon humanity and that the noble savage had lived much better without them.

Today we tend to think of agriculture as an unmitigated good. Look around, we might say, and imagine all *this*—gesturing expansively—without the grains that allow millions of people simply to eat. But our belief that agriculture in the present day is a blessing to humankind, which undoubtedly it is, does not commit us to insisting that agriculture, in all its forms, in all times and places, has always been a boon to everyone. Nor does the view that agriculture began as a curse commit us to believing that agriculture remains a curse today. Reality is allowed to be complex like that, even where a sentimental ideology like Rousseau's cannot be.

At all stages of historical and prehistorical development, critical evaluation ought to be key. Not at all trivially, we might ask what blessings agriculture brought to the slaves of recorded history. "They get to eat" appears grossly unsatisfying given all the other ways that agriculture imposed itself on them. For free people, though, "they get to eat" isn't so bad. With this in mind, the conflicts between ideological priors and archaeological discoveries may just be mitigated. This new story of the past in no way precludes libertarianism.

It bears noting that Scott is a writer of extraordinary talent. His prose is gorgeous, and annoyances are few. One of them, though, is as follows: he constantly peppers his work with anachronistic terms—"capitalism," "proletariat," "gulag," even "human resources." At times these anachronisms both entertain and instruct: war captives in Mesopotamia clearly fit Marx's characterization of the proletariat rather well, even if they did not toil at sophisticated machines owned

by others. They were alienated from their labor rather more directly because they were enslaved.

Those early slaves may be much more of a proletariat than the industrial workers of the 19th century ever were. Industrial workers' living conditions have steadily improved, and the lines between them and the other social classes have blurred rather than sharpened. War captives in early civilizations faced a different fate, and it is both just and ironic that Scott deploys the word "gulag" to name the detention camps in which they were placed.

At other times, though, the effort to connect the ancient world to the modern seems contrived, as when slavery is characterized as a "human resources" strategy. Perhaps in *your* human resources department, I am tempted to answer. The functionalist observation that both are about the supply of labor seems insufficient to save the analogy. Readers would do well to interrogate these terms as much for what they fail to say as for what they succeed at saying. In both cases they are instructive, and the constant interplay between the present and the distant past is one of the most appealing aspects of this book. To what extent are we still the kept creatures of the state and of the agriculture that the state imposed on us? Must that system, or any part of it, *always* be a curse? It will always be profitable to reflect on questions like those.

Jason Kuznicki
Cato Institute

DISCRIMINATION
and
DISPARITIES

Thomas Sowell

BASIC BOOKS \$28.00 179 PAGES

Discrimination and Disparities challenges believers in such one-factor explanations of economic outcome differences as discrimination, exploitation or genetics. It offers its own new analysis, based on an entirely different approach— and backed up with empirical evidence from around the world.

The point is not to recommend some particular policy “fix,” but to clarify why so many policy fixes have turned out to be counterproductive, and to expose some seemingly invincible fallacies behind many of those counterproductive policies.

The multiple, and sometimes mutually contradictory, definitions of “discrimination” are sorted out, along with the varying costs of discrimination— to the victims, to society and even to the discriminators themselves— and the consequences of those costs in affecting behavior in different economic institutions.

The role of chance and the pitfalls of statistics, even when the numbers themselves are accurate, are pointed out as factors in misunderstandings of the causes of disparities in such things as income, education and outcomes in the criminal justice system. This is a book not only about economics but also about public policy and social issues.

Thomas Sowell is a Senior Fellow at the Hoover Institution, Stanford University. He has taught economics at Cornell, UCLA, Amherst and other academic institutions. His books have been on subjects ranging from economics to philosophy, history and decision-making theory, and his articles have appeared in both scholarly journals and in such popular media as the *Wall Street Journal* and *Fortune* and *Forbes* magazines.

Cato Institute

Founded in 1977, the Cato Institute is a public policy research foundation dedicated to broadening the parameters of policy debate to allow consideration of more options that are consistent with the principles of limited government, individual liberty, and peace. To that end, the Institute strives to achieve greater involvement of the intelligent, concerned lay public in questions of policy and the proper role of government.

The Institute is named for Cato's Letters, libertarian pamphlets that were widely read in the American Colonies in the early 18th century and played a major role in laying the philosophical foundation for the American Revolution.

Despite the achievement of the nation's Founders, today virtually no aspect of life is free from government encroachment. A pervasive intolerance for individual rights is shown by government's arbitrary intrusions into private economic transactions and its disregard for civil liberties. And while freedom around the globe has notably increased in the past several decades, many countries have moved in the opposite direction, and most governments still do not respect or safeguard the wide range of civil and economic liberties.

To address those issues, the Cato Institute undertakes an extensive publications program on the complete spectrum of policy issues. Books, monographs, and shorter studies are commissioned to examine the federal budget, Social Security, regulation, military spending, international trade, and myriad other issues. Major policy conferences are held throughout the year, from which papers are published thrice yearly in the *Cato Journal*. The Institute also publishes the quarterly magazine *Regulation*.

In order to maintain its independence, the Cato Institute accepts no government funding. Contributions are received from foundations, corporations, and individuals, and other revenue is generated from the sale of publications. The Institute is a nonprofit, tax-exempt, educational foundation under Section 501(c)3 of the Internal Revenue Code.

CATO INSTITUTE
1000 Massachusetts Ave., N.W.
Washington, D.C. 20001
www.cato.org