Trust Doesn't Explain Regional U.S. Economic Development and Five Other Theoretical and Empirical Problems with the Trust Literature

By Alex Nowrasteh and Andrew C. Forrester

October 8, 2019

CATO WORKING PAPER
No. 57

CATO INSTITUTE
1000 Massachusetts Avenue NW
Washington DC 20001

Cato Working Papers are intended to circulate research in progress for comment and discussion.
Available at www.cato.org/workingpapers.

Trust Doesn't Explain Regional U.S. Economic Development and Five Other Theoretical and Empirical Problems with the Trust Literature*

Alex Nowrasteh
Cato Institute

Andrew C. Forrester
Cato Institute

This version: October 2019

Abstract

Economists have developed a vast empirical literature on how cultural traits like generalized trust affect economic output. Much of this literature finds a positive causal relationship between measures of generalized trust, as gathered by international surveys, and economic output. However, the trust literature commits five deadly empirical and theoretical sins that undermine its findings. From the quality of the survey questions and responses to the paucity of theoretical models used to explain how trust affects economic outcomes to the radically different results from experimental evidence, the trust literature is riven with poor methods and bad data that undermine its conclusions. Even so, applying the best methods in the trust literature to regional level analysis in the United States reveals no statistically significant correlation between economic output and trust. We see no reason to trust the findings of the trust literature.

JEL Codes: B4, D7, O5

Key Words: Culture, Trust, Institutions

*Alex Nowrasteh (Corresponding author) is Director of Immigration Studies at the Cato Institute, his email is anowrasteh@cato.org; Andrew C. Forrester is a Research Associate at the Cato Institute, his email is aforrester@cato.org.

1 Introduction

Culture is a vast concept, but most cultural traits such as food or clothing styles have zero impact on economic growth. Economists thus face a challenge when incorporating culture into economic models and identifying which portions of it affect economic behavior. Economists usually begin by treating culture as a black box whereby inputs enter and outputs leave after being transformed in the dark. As the black box metaphor demonstrates, economists make few attempts to explain how those inputs are transformed into outputs via culture. To fill the culture black box, some economists concentrate on measures of generalized trust (henceforth trust) as a proxy measure for economically-relevant culture (Gambetta, 2000).

Trust is the "the subjective probability with which an agent assesses that another agent or group of agents will perform a particular action" (Gambetta, 2000). Economists have settled on trust as an economically relevant cultural trait for two main reasons. The first is that trust seems like it can be incorporated into standard economic models, although it rarely is (Algan and Cahuc, 2010, 2013; Guiso, Sapienza and Zingales, 2009b; Zak and Knack, 2001). The second is data availability (Weil, 2005). Surveys like the World Values Survey (WVS), EuroBarometer, the American General Social Survey (GSS), the Latinobarómetro, and others have all asked similar questions about trust for decades in many different countries. The specific trust question used in those surveys asks respondents: "Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?" The responses are "Most people can be trusted," "Can't be too careful," and "Depends" (Algan and Cahuc, 2010). Economists interpret the response "Can't be too careful" as low trust.

There are few theoretical microeconomic models for how trust or other social norms could affect production on the firm level, but none for the macroeconomic level. One informal model assumes that transaction costs are higher when trust is lower, which diminishes growth. Another informal model is simply that high trust reduces the resources that firms and individuals rationally spend on protective purposes (Bjørnskov, 2018). Another is a more formal model of innovative investment that includes social norms, social trust, and networks with reciprocity to model investment in productive ideas that have economy-wide effects (Akcomak and Weel, 2009; Ikeda, 2008). A last formal micro model embeds a principal-agent situation in a model of education and monitoring costs (Bjørnskov, 2018). None of these formal models link their microlevel effects onto a macroeconomic growth model (Rose, 2011, p. 171). Thus, the "empirical literature has proceeded ... without much clear interaction with theoretical development" (Bjørnskov, 2018).

2 Five Deadly Sins of the Trust Literature

The trust literature is vast and covers many different countries. However, there are serious methodological and data problems that cast significant doubts on empirical trust-related findings in the economics literature. The broad problems with the trust-growth literature are that there are no macroeconomic growth models with microeconomic foundations based on trust,[1] the trust survey questions are internally invalid, and the methodology for making sense of those weak measurements is so poor that we cannot currently identify a correlative or causal link between trust and any indicator of economic output. In other words, the vast trust-growth literature is empirically suspect as "too many theoretical and empirical problems are associated with this measure and the theoretical construct to claim that the trust literature has showed that culture affects economic growth" (Beugelsdijk and Maseland, 2011, p. 222) We call these specific problems the five deadly empirical and theoretical sins of the trust literature.

The first deadly sin of the trust-growth literature is that it contains no macroeconomic growth model that incorporates trust, either in its micro-foundations or otherwise (Beugelsdijk and Maseland, 2011, p. 213). Furthermore, the trust-growth literature does not contain a formal theory of social capital formation broadly or one of trust specifically (Guiso, Sapienza and Zingales, 2011, p. 469). Most trust researchers aggregate assumed efficiencies at the microeconomic level up to the macroeconomic level and assume that trust creates economy-wide growth: an illegitimate leap in the logic of micro- to- macro functioning (Beugelsdijk and Maseland, 2011, p. 208). The relationship between an individual's trust and income may not be true for society and cannot be aggregated up to form a truthful representation of the whole (Beugelsdijk and Maseland, 2011, p. 208).

Economists working on trust have mostly skipped the model-creation phase and focused

[1] Weil (2005) is the closest.

on empirical testing after assuming that such a relationship exists. This approach has produced predictably meaningless empirical results due to the lack of a rigorous theoretical foundation (Herrmann-Pilath, 2010; Kapas, 2017). One partial exception to this is by Butler, Giuliano and Guiso (2016) who create a micro-level model to explain individual investments in trust and how it relates to personal income. Testing his model against individual trust survey data from Europe shows a distinct n-shaped relationship where too little trust and too much trust is correlated with lower individual income (Butler et al., 2016). However, Butler et al. (2016) do not even attempt to build their microeconomic model into a macroeconomic growth model and avoid making any claims about trust's impact on national level output.

The second deadly sin of the trust literature is that the trust question itself does not produce internally valid responses. Recall, the trust question is: "Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?" There is no universal measure of trust because, in part, its meaning is culturally and contextually specific (Beugelsdijk and Maseland, 2011, p. xvii). The responses to the trust question are "Most people can be trusted," "Can't be too careful," and "Depends." The meaning of those responses is also unclear (Moore, 1999), as Putnam (2000) explains:

> This question clearly taps feelings about the trustworthiness of the generalized other – thin trust – but the meaning of the responses remains murky in one respect. If fewer survey respondents nowadays say, "Most people can be trusted" that might mean any one of three things: 1) the respondents are accurately reporting that honesty is rarer these days; or 2) other people's behavior hasn't really changed, but we have become more paranoid; or 3) neither our ethical demands nor other people's behavior have actually changed, but we now have more information about their treachery, perhaps because of more lurid media reports (Putnam, 2000, pp. 137–138).

Putnam (2000) questioned the meaning of the "Most people can be trusted" response to the trust question, but Miller and Mitamura (2003) questioned the meaning of the "Can't be too careful" response. Miller and Mitamura (2003) argue that the "Can't be too careful" response measures respondents' levels of caution, not distrust. When they replace the cautionary "Can't be too careful," response with another response that indicates distrust on a separate scale, Americans respondents turn out to have higher levels of trust than do Japanese respondents even though Japanese are famously far more trusting than Americans in the standard WVS version of the question (Beugelsdijk and Maseland, 2011, p. 216). Societies where different degrees of caution are used, independent of the idea of trust, will bias responses to the trust question which means that the responses are likely incomparable between countries. For instance, where people live in closed and safe environments with low levels of caution, responses to the trust question may overestimate true trust (Beugelsdijk and Maseland, 2011, p. 217).

There are other problems with the internal validity of the trust question. Glaeser, Laibson, Scheinkman and Soutter (2000) argue that the trust question measures respondent opinions about the trustworthiness of others and not their personal level of trust itself. When faced with the risk of other people being untrustworthy, what matters most is one's aversion to betrayal which is not correlated with trust in cross-country analyses (Bohnet, Greig, Herrmann and Zeckhauser, 2008). In one famous experiment, social scientists dropped thousands of wallets with $50 cash and contact information in countries around the world to look at the return rates and found a 0.67 correlation between wallet return rates and perceptions of generalized trust (Felten, 2000; Weil, 2005). On the other hand, different studies show that residents of diverse neighborhoods in Chicago with low levels of trust are no less likely to return lost letters or administer CPR to strangers than those from high trust neighborhoods (Abascal and Baldassarri, 2015; Iwashyna, Christakis and Becker, 1999; Sampson, 2012).

Respondents also understand the question differently. There is also ambiguity about what the "most people" portion of the trust question means, as many respondents think of social circles that they rarely interact with when picturing "most people" even though they trust familiar social circles that they frequently interact with (Delhey and Newton, 2005; Hardin, 2002). Responses to the trust question also depend on when the question is asked in the survey (T. Smith, 1997; Uslaner, 2002). Additionally, local variation in cultural meanings of trust and in respondents' socioeconomic characteristics also affect responses (Roth, 2009, ft. 3).

The trust question is also not internally valid over time within countries. For example, Wave 5 of the WVS shows that Iran has a very high average trust score of 65.3 percent that drops to a low 10.6 in Wave 6. Similar discrepancies persist across multiple questions unrelated to trust, such as happiness and belief in environmental causes, suggesting that the WVS is not useful in making comparisons over time even in the same country.

Müller, Torgler and Uslaner (2012) also note major inconsistencies between the WVS estimates of trust relative to other comparable surveys, particularly the sizable systematic differences between specific waves of the WVS relative to comparable national surveys conducted at the same time in Canada, China, Indonesia, Iran, the United Kingdom, Vietnam, and other countries. Figure 1 visualizes the differences between responses to the identically worded trust question in the American GSS and WVS over time. Matching each WVS sample to corresponding years in the GSS, we compute weighted average trust scores for each time interval. While GSS and WVS responses overlap in many periods, we do find a large difference of nearly 14 percentage points in generalized trust over the 1990-1993 period, a 7.2 percentage point difference over the 2005-2009 period, and a 3.6 percentage point difference over the 2010-2014 period. Out of six possible time series data points, we find that American responses to the WVS trust question only match 50 percent of responses to the American GSS. These problems point to a major and irregular discrepancy in responses to the same question asked in the same years in the same country by different sources.

There is a larger literature of trust on the organizational level, but the measurements are so different from those on the national level and in the experimental economics literature that they are difficult to compare (Beugelsdijk, 2006). For instance, organizations-level trust questions measure trust between specific partners or individuals in an organization. They also measure trust on a 1-5 scale or a 1-7 experience scale, as related to the proportion of promises kept, the provision of trustworthy information, and inter-firm relations (Beugelsdijk and Maseland, 2011, p. 218). There is little compelling evidence of any relationship between individual level responses in organizational surveys to the trust question and nationwide responses.

The WVS measure of trust lacks comparability across time, space, and with other surveys that use identical wording for their trust question. The combination of these problems undermines the validity of the WVS-based measure itself and other national measures of trust that place nearly insurmountable data constraints on further empirical research regarding trust (Beugelsdijk, de Groot H.L.F. and van Schaik, 2004).

The third deadly sin of the trust literature is that responses to the trust question do not generally predict trusting behavior in real-world micro-level experiments or in trust games (Algan and Cahuc, 2013; Ermisch, Gambetta, Laurie, Siedler and Uhrig, 2009; Glaeser et al., 2000; Glaeser, Liabson and Sacerdote, 2002; Karlan, 2005; Näf and Schupp, 2009; Rose, 2011; Sapienza, Toldra-Simats and Zingales, 2013), with some occasional exceptions (Fehr, Fischbach, Rosenbladt, Schupp and Wagener, 2002; Felten, 2000).

The typical trust game works like this: The game has two players, and each make a single choice. They usually can not see each other, and they typically do not know who they are playing as they are often behind a computer. For the first move, Player 1 starts with $5 and decides how much money to send to Player 2 and how much to keep for himself, if any. Money sent over by Player 1 triples in value, so if Player 1 sent over $5, Player 2 now has $15. Player 2 then decides how much money to send back, if any. Most players return about the amount that Player 1 initially sent over (Jones, 2016). Although many researchers cannot agree on how to interpret the results from trust games, the actions of Player 1 are supposed to show his level of trust while Player 2's actions are supposed to measure his level of *trustworthiness* (Glaeser et al., 2002; Sapienza et al., 2013).

People are always more trusting and trustworthy in experiments and games than they would appear to be in surveys (Rose, 2019). Seventy-four percent of Swedes and 41 percent of Tanzanians say that most people can be trusted on surveys, but their actual behavior in the trust game is remarkably similar to each other (Beugelsdijk and Maseland, 2011; Holm and Danielson, 2005). American surveys consistently show that blacks report having less trust than whites, but blacks are as trusting and more trustworthy than whites in trust games even though both groups trust members of their own race more than members of the opposite race. As it turns out, responses to the trust question are not predictive of black behavior in the trust game, but they do predict white behavior (Abascal and Baldassarri, 2015; Simpson, McGrimmon and Irwin, 2008).

The ultimatum game is another laboratory experiment that is supposed to measure trust. The typical ultimatum game works like this: The game has two players, and each make a single choice. Player 1 proposes how to divide up a pot of money with Player 2. If Player 2 accepts Player 1's division, then the money is divided up thusly. If Player 2 rejects the proposed division, neither player gets any money. Again, there is no statistical relationship between individual actions in the ultimatum game and their responses to WVS trust questions (Fernandez, 2011; Oosterbeek, Sloof and van de Kuilen, 2004).

It is also important to move beyond the laboratory. Micro-level experiments in Peruvian villages with micro loans show that responses to the trust question successfully predict default on a loan and other trustworthy actions, but they fail to predict savings or trusting actions (Karlan, 2005). The underlying non-formalized model for why trust matters is that countries "characterized by trust are economically more successful is mostly based on micro-level arguments deduced from either transaction costs theory or game theory ... [and] reduced transactions costs and principal-agent problems ... increase (the efficiency of) investments in physical and human capital, and promote innovation" (Beugelsdijk and Maseland, 2011, p. 208). The divergence between actions in laboratory and micro-level experiments and responses to trust survey questions raises significant problems with that underlying non-formalized model. The surveys could be flawed, the experiments could be flawed, or both could be flawed, but the radically different results show that at least one of those is true and calls into question the trust data.

The fourth deadly sin is that many of the major papers in the trust literature are contaminated by various types of sample biases. The first is country sample selection bias whereby results change significantly based on the sample of countries chosen (Durlauf, 2002). The canonical paper by Knack and Keefer (1997) chose 27 OECD countries and 2 non-OECD countries in their sample: India and Nigeria. Zak and Knack (2001) use an overlapping sample of 42 countries. These sample selection problems are compounded because the WVS does not have consistent sampling methods across countries and time. The second type of sample selection bias occurs when researchers compare the trust level of immigrants in destination countries with the trust level of their former co-nationals in their homelands (Algan and Cahuc, 2010). Emigrants likely have different levels of trust than the non-emigrant citizens from their home countries (Uslaner, 2008). Additionally, emigrant self-selection likely affects measurements of trust in later generations and is even further disturbed by ethnic attrition or the selective and biased self-identification of the descendants of immigrants with their ethnic and racial groups (Duncan and Trejo, 2016; Fernandez, 2011; Giavazzi, Petkov and Schiantarelli, 2019; Weil, 2005).

Another source of sample selection bias is that many of the findings in this literature depend on the years chosen for study. For instance, Algan and Cahuc (2010) use survey responses from elderly immigrants decades after they arrived in the United States to predict the levels of trust in their home countries generations in the past. Müller et al. (2012) show that the reported amount of trust-persistence over generations in Algan and Cahuc (2010) is not robust in different waves of the same survey, which raises serious doubts about the robustness of their results. For instance, Müller et al. (2012) note that WVS Wave 4 for the years 1999-2004 has significantly different trust responses relative to many other survey datasets to such an extent that the results presented in Algan and Cahuc (2010) were not robust in other waves of the WVS (Clemens and Pritchett, 2019). This lack of replicability for other waves suggests that the results in Algan and Cahuc (2010) may merely be an artifact of WVS mismeasurement in a single wave which casts significant doubt on their trust estimates decades in the past.

The fifth deadly sin is that even if the trust question were free from measurement error or sample selection bias, trust may be a proxy measurement for other deeper causes of economic development (Clemens and Pritchett, 2019). Omitted variable bias and endogeneity are persistent problems in this literature that have not been satisfactorily resolved. Any feature of a country that lastingly affects both trust and development for more than four generations could generate correlation in the absence of causation, such as slavery affecting levels of trust in Africa and among the descendants of African slaves in the United States today (Nunn, 2008; Nunn and Wantchekon, 2011). Thus, the Algan and Cahuc (2010) methods would generate correlation between inherited trust in the United States and contemporary economic outcomes in the countries of origin that do not arise causally from trust's impact on economic outcomes (Clemens and Pritchett, 2019).

Papers in the trust literature often estimate cross-country Barro-style growth regressions using cross-sectional datasets or panel data models (Blume, Brock, Durlauf and Ioannides, 2010; Knack and Keefer, 1997; Roth, 2009). A critical consideration is whether the parameters underlying these econometric models are, under a set of reasonable assumptions, truly identified (Blume et al., 2010; Durlauf, 2002). Consider the basic Barro-style growth regression framework from Bazzi and Clemens (2013):

$$g = \alpha + \gamma T + \sum_{k=1}^{K} \beta_k x_k + \epsilon, \qquad (1)$$

where g is economic growth; T is a measure of generalized trust; $\{x_k\}$ is a set of K, likely endogenous, regressors; and ϵ is a disturbance term. Bazzi and Clemens (2013), Durlauf (2002) describe the precarious nature of searching for viable instruments in cross-country growth regressions. Suppose that we observe some instrument z that satisfies the relevance and exclusion conditions that $cov(T, z) \neq 0$ and that $E(z\epsilon) = 0$. For the instrument z to pass the exclusion restriction, it must also be that $cov(x_k, z) = 0 \forall k$. In other words, the instrument must explain only variation in trust while being uncorrelated with other relevant growth determinants x_k to be a valid instrument. Insofar as there is correlation between a candidate instrument for trust and other determinants of growth, the coefficient describing the relationship between trust and growth will not be identified (Bazzi and Clemens, 2013; Durlauf, 2002). There are no viable instruments yet identified in this literature and it is unlikely that any will be (Durlauf, 2002; Guiso et al., 2011). Despite that, economists have chosen literacy rates (Tabellini, 2010), historical political institutions (Akcomak and Weel, 2009; Tabellini, 2008, 2010), ethno-linguistic fractionalization (Knack and Keefer, 1997), trust levels in immigrant home countries (Algan and Cahuc, 2010), and common religion (Guiso, Sapienza and Zingales, 2009a) as instruments even though none of them are "exogenous to the error term beyond doubt" (Fehr, 2009). Growth studies that use instrumental variable strategies must be able to demonstrate an instrument's strength and its validity even though theories of economic growth are not inherently mutually exclusive (Bazzi and Clemens, 2013; Durlauf, 2002; Durlauf and Quah, 1999).

Instruments may pass quantitative tests of overidentification and weak instrumentation but establishing their validity with respect to the structural growth equation requires additional, rigorous theoretical justifications to pass the exclusion restriction (Durlauf, 2002; Guiso et al., 2011). None of the instruments listed above satisfy those additional rigorous justifications in the trust literature (Durlauf, Johnson and Temple, 2005). A prime example of this is the use of instruments based on the history of local political independence to causally measure the impact of civic capital and trust on per capita income in towns in Northern Italy (Guiso et al., 2011, 2016). Since the cities in Northern Italy have been governed by the Italian state since the 19th Century, the differences in culture across regions can theoretically explain the various economic outcomes and be instrumented for with a measure of historical independence. However, the historical instrument does not work for measuring the causal impact of social capital or trust on any economic outcome:

> For the instruments to be valid, it must be that the historical episodes that built up civic capital did not at the same time foster the accumulation of other forms of capital that have lasted to today and still exert a direct influence on income. For instance, in the Guiso et al. (2009a) context, having been a free city in the 13th century may have resulted in accumulated assets of some sort that still directly affect income today, besides affecting it indirectly because of its boost on civic capital. Using the Bishop city and the Etruscan city indicators, which proved to be good instruments for the historical determinants of civic capital, is not a solution either. In fact, even if they affect civic capital only because they facilitated the emergence of the free city (and thus qualify as instruments in a civic capital regression), they also boosted all the unobservable assets that may continue to affect a city's income today (which may invalidate them as instruments in an income regression). The only way to account for this is to obtain direct measures of these assets and try to control for them. The general point is that historical shocks to civic capital could have also shocked other types of capital that are as persistent as civic capital and which may have an independent, direct effect on income (Guiso et al., 2011).

Most papers in the trust literature rely on estimating multiple regression models, either in the form of cross-sectional or panel data models (Knack and Keefer, 1997). However, the parameters underlying the econometric models in the trust literature are not truly identified under any reasonable set of yet-considered assumptions (Blume et al., 2010; Durlauf, 2002).

The sixth deadly sin of the empirical trust literature is that sub-national level data in the United States that is collected under better conditions do not indicate a robust positive relationship between trust and growth. As noted above, it is difficult to compare survey responses to the trust question between countries for many reasons (Durlauf, 2002). Subnational survey response data are less susceptible to significant cross-cultural differences in responses but still commit some of the other deadly sins noted in this section.

The next section builds the best possible empirical case to test the relationship between trust and growth inside of the United States through subnational response to the GSS. Even under this better condition inside of one country, the results show no statistically significant relationship between trust and growth inside of the United States while being almost as statistically problematic as the results in the rest of the trust literature.

3 Trust and Income in American Geographical Regions

Many of the data and survey problems described above are largest for trust comparisons between countries. Internal survey validity concerns and cross-sectional limitations are insurmountable in cross-country regression analyses, but large-scale social surveys inside of countries could provide a better source of data for descriptive time-series analysis. Although it is not ideal, the American GSS asks a trust question and allows a subnational level of geographic aggregation inside of the United States (T. W. Smith, Davern, Freese and Morgan, 2018). Dincer and Uslaner (2010) exploit state-level data and argue that focusing on the effect of regional variation in trust on regional growth inside of a country removes many concerns over internal validity of the survey question controls for some unobservable differences across countries that are difficult to control for in cross-sections. Although they find a robust relationship between state-level economic growth and trust inside of the United States, they rely on expensive to acquire state-level GSS data pooled into two cross-sections for the 1990s, compare state-level GDP growth without controlling for population, and they assume homoskedasticity in reporting their first stage F statistic which makes it unclear whether their instrument is valid because .

We thus undertake a regional level analysis of how trust is related to per capita personal income by region inside of the United States. Regions of the United States provide better data to study the impact of trust on growth than Europe does because regional trust variation is greater inside of the former. In empirical studies that examine European-regional trust variation, almost all countries have near-uniform levels of trust across their subnational regions. Thus, empirical identification for European-regional regressions will only result from those countries with substantial regional variations in trust – primarily Belgium, Italy, and Spain. Using an estimator with country fixed effects or trust measures relative to the country average produces almost no systematic regional trust variation from relatively trust-homogeneous countries like Denmark or the Netherlands (Bjørnskov, 2018, p. 547). The United States suffers less from that European-wide problem because of its greater internal variation on responses to the trust question.

Despite the empirical limitations in regional-level analyses of trust's impact on growth, they are probably the best way forward (Dincer and Uslaner, 2010). The smallest level of geography available in the GSS public files is the Census Division, which represents nine consistent groups of three to eight states by region of the United States. The limited number of geographic subunits, the lack of a link between behavior in games and experiments with answers to the trust question, and the likely impossibility of identifying a valid instrument mean that our results will be up to the standards of the trust literature, but those standards are low (Blume et al., 2010; Durlauf, 2002). Regardless, this attempted analysis of the subnational GSS data is valuable because it demonstrates the weakness of the empirical trust literature in the best-case scenario that compensates for some of the other deficiencies in regional-level analyses (Dincer and Uslaner, 2010).

This section relies on a panel of the nine regional Census Divisions over the span of 1972-2016. Beginning in 1972, the GSS collected data on an annual basis over the span of 1972-1994 (missing the years 1979, 1981, and 1992) and biennially in even years from 1994 through 2018. From the GSS, we measure trust as the weighted average positive responses to the trust question in each region.[2] To arrive at complete time series for the growth measure, we linearly interpolate any gaps in the trust-response time series. Data on per capita personal income comes from the Bureau of Economic Analysis (BEA) Regional Economic Accounts that we aggregated up to the Census Division level using the Census Bureau's state-to-division contingency tables. Finally, we adjust the annual per capita personal income data for inflation using the personal consumption expenditures (PCE) price index.

We use Barro style growth regressions that are prominent in the existing trust and growth literature (Roth, 2009). The main estimating equation is the standard additive fixed-effects model:

[2]We use the person level `WTSALL` weights.

$$g_{i,t} + \alpha_i + \lambda_t + \gamma T_{i,t-1} + \epsilon_{i,t}, \tag{2}$$

where $g_{i,t}$ is the growth (log difference) in per capita income for Census Division i in year t, $T_{i,t-1}$ denotes the average generalized trust level in year $t-1$, α_i and λ_t are Census Division and year fixed effects, and $\epsilon_{i,t}$ is a distrubance term. Following Roth (2009), each variable is lagged behind one period to reduce endogeneity concerns with respect to contemporaneous reverse causality between trust and growth. We express generalized trust as a natural log, thereby making the coefficient of interest γ an elasticity between average trust on growth in per capita income.

Since the estimate $\hat{\gamma}$ is unbiased when lagged trust $T_{i,t-1}$ is orthogonal to the error term $\epsilon_{i,t}$, (2) provides only a biased measure of correlation between lagged trust and growth. *These regressions should therefore be interpreted as only correlative and subject to significant bias due to measurement error and endogeneity.* Recall that in the additive fixed-effects specification, the bias of the OLS estimator $\hat{\gamma}$ will be biased towards zero in the presence of an additive measurement error in $T_{i,t-1}$.

Another key consideration pertains to the limited data available in the GSS public use files. In particular, the public-use GSS files only contain data on the 9 Census regions. To correct for serial correlation in the residuals, one would typically use a clustered robust variance estimator (CRVE) to compute standard errors; however, the CRVE is notably biased in the presence of very few clusters. In particular, Cameron, Gelbach and Miller (2008) note that the standard CRVE tends to over-reject in the presence of few clusters. To ensure proper inference, we use a wild cluster bootstrap (WCB) to compute p-values and confidence intervals for our point estimates in line with Cameron et al. (2008) and Roodman, Nielsen, MacKinnon and Webb (2019).

The West North Central Division contains Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, and South Dakota. The East South Central Division contains Alabama, Kentucky, Mississippi, and Tennessee. The least volatile Census Divisions with respect to responses to the trust question are the Middle Atlantic Division and the Pacific Division. The Middle Atlantic divisions contains the states of New Jersey, New York, and Pennsylvania while the

Pacific Division contains the states of Alaska, California, Hawaii, Oregon, and Washington.

Table 1 shows the summary statistics of each time-series of trust scores by Census Division. Trust scores are more volatile in earlier years of the GSS and often vary by 10 percentage points from one year to the next for many Census Divisions. The West North Central, East South Central, and Mountain Divisions are most volatile over time with the highest standard deviations in the sample (Table 1). Figure 2 presents time-series plots of trust for each Census Division. Like the cross-country analyses of responses to the trust question, sub-national data from the United States Census Divisions show a wide degree of variance over time and between regions.

In line with other studies in the trust literature, we next test the correlation between average trust and economic growth in the multiple regression framework introduced above. It is important to repeat that any results here are purely correlative and not causal. We are committing some of the five deadly empirical sins with this regression to demonstrate the weakness of the empirical trust literature in this best-case scenario. Regardless of our sins, this attempted analysis of the subnational GSS data is valuable to demonstrate the weakness of the empirical trust literature in the best-case scenario because the results do not show a link between trust and growth.

Data limitations in the GSS public-use data relegate us to a small cross-sectional dimension of the nine Census Divisions, which limits our ability to provide reasonable estimates of variance for parameter estimates. The unbiasedness of the OLS estimator in the face of endogeneity means that we cannot provide accurate estimates for any relationship that emerges, which is a ubiquitous issue in the rest of the trust and growth literature that few researchers confront.

The results from the fixed effects regressions of log-differenced per capita income growth on lagged trust and log income are presented in Table 2. We consider two estimation samples based on the different frequencies of GSS observations – annual observations and biennial observations. We also partition unobserved heterogeneity into two functional forms: additive Census Division and year fixed effects and Census Division-specific linear time trends.

Our main results for the biennial GSS data are shown in Table 2. Columns 1 and 2 report estimated elasticities between generalized trust and personal income growth and include year and division-specific linear time trends, respectively. Following Cameron et al. (2008) we report wild cluster bootstrap (WCB) p-values in parentheses and confidence intervals in brackets. In each case, we find that the coefficient estimates for generalized trust are very close to zero and statistically insignificant. Similar to other studies, such as Roth (2009), we also account for persistence in growth by including lagged per capita income. Results from these regressions in columns 3 and 4 show a similar pattern of insignificance. In both specifications, we find that trust is statistically insignificant after accounting for persistence for both year fixed effects and division-specific trends.

To exploit year-to-year variation in per capita income growth, we further examine annual data on growth and trust in Table 3. Again, we find that trust is statistically indistinguishable from zero after accounting only for fixed effects and linear trends in columns 1 and 2. This null result persists after controlling for persistence in columns 3 and 4, for which we again find no significant statistical relationship between trust and growth.

Measurement error in the trust literature is widely documented. Biases force the estimates $\hat{\gamma}$ in each regression specification to zero, as evident by the very small magnitude of each estimated parameter, which tentatively suggests that statistical biases from endogeneity or measurement error are a major influence on the estimated relationship between trust and growth (Table 2). There are no valid instruments that can resolve the endogeneity problem. Combining these results, we find no evidence of a relationship between trust and economic growth using the GSS panel of subnational data.

4 Conclusion

Contributions to the trust literature have yet to overcome "serious measurement error" and the methodological problems explained above and demonstrated in our subnational examination of the nine Census Divisions in the United States (Blume et al., 2010). Across the many dimensions of analysis available to researchers, ranging from cross-sections to panel data, the trust studies provide neither compelling nor empirically valid identification strategies to derive meaningful statistical relationships between trust and growth (Blume et al., 2010; Durlauf, 2002). The lack of theoretical guidance underlying various studies' econometric identification strategies cast doubt on the validity of their estimates (Durlauf, 2002). The lack of statistical rigor means that economists are not measuring what they think they are measuring when they regress trust on an economic indicator (Beugelsdijk, 2006; Kapas, 2017). As long ago as 2002, Durlauf (2002) argued that the usage of observational survey data is likely to be less compelling relative to the experimental evidence that, subsequently, has shown little relationship between responses to the trust survey question and behavior in the laboratory. It is a shame that so few heeded Durlauf's (2002) advice and spent so much time attempting to squeeze statistical insight from poorly measured survey noise (Blume et al., 2010; Durlauf, 2002).

References

Abascal, M. and Baldassarri, D. (2015). *Love thy neighbor? Ethnoracial diversity and trust reexamined*. *American Journal of Sociology*, *121*(3), 722–82.

Akcomak, S. and Weel, B. (2009). *Social capital, innovation and growth: Evidence from Europe*. *European Economic Review*, *53*(5), 544–67.

Algan, Y. and Cahuc, P. (2010). *Inherited trust and growth*. *American Economic Review*, *100*(5), 2060–92.

Algan, Y. and Cahuc, P. (2013). *Trust and growth*. *Annual Review of Economics*, *5*(1), 521–49.

Bazzi, S. and Clemens, M. (2013). *Blunt instruments: Avoiding common pitfalls in identifying the causes of economic growth*. *American Economic Journal: Macroeconomics*, *5*(2), 152–86.

Beugelsdijk, S. (2006). *A note to the theory and measurement of trust in explaining differences in economic growth.* *Cambridge Journal of Economics*, *30*(3), 371–87.

Beugelsdijk, S., de Groot H.L.F. and van Schaik, T. (2004). *Trust and economic growth: A robustness analysis*. *Oxford Economic Papers*, *56*(1), 118–134.

Beugelsdijk, S. and Maseland, R. (2011). *Culture in economics: History, methodological reflections, and contemporary applications*.

Bjørnskov, C. (2018). *Social trust and economic growth*. In E. Uslaner (Ed.), *The oxford handbook of social and political trust.*

Blume, L. E., Brock, W. A., Durlauf, S. N. and Ioannides, Y. M. (2010). *Identification of Social Interactions* (Economics Series No. 260). Institute for Advanced Studies. Retrieved from https://ideas.repec.org/p/ihs/ihsesp/260.html

Bohnet, I., Greig, F., Herrmann, B. and Zeckhauser, R. (2008). *Betrayal aversion: Evidence from Brazil, China, Oman, Switzerland, Turkey, and the United States*. *American Economic Review*, *98*(1), 294–310.

Butler, J., Giuliano, P. and Guiso, L. (2016). *The right amount of trust*. *Journal of the European Economic Association*, *14*(5), 1155–80.

Cameron, A., Gelbach, J. and Miller, D. (2008). *Bootstrap-Based Improvements for Inference with Clustered Errors*. *Review of Economics and Statistics*, *90*(3), 414–27.

Clemens, M. and Pritchett, L. (2019). *The new economic case for migration restrictions: An assessment*. *Journal of Development Economics*, *138*, 153–64.

Delhey, J. and Newton, K. (2005). *Predicting cross-national levels of social trust: Global pattern or Nordic exceptionalism?* *European Sociological Review*, *21*(4), 311–27.

Dincer, O. and Uslaner, E. (2010). *Trust and growth*. *Public Choice*, *142*(59), 60–61.

Duncan, B. and Trejo, S. (2016). *The complexity of immigrant generations: Implications for assessing the socioeconomic integration of Hispanics and Asians*. *Industrial and Labor Relations Review*, *70*(5), 1146–75.

Durlauf, S. (2002). *On the empirics of social capital*. *The Economic Journal*, *112*(483), 459–79.

Durlauf, S., Johnson, P. and Temple, J. (2005). *Growth econometrics*. In S. Durlauf and P. Aghion (Eds.), *Handbook of economic growth*.

Durlauf, S. and Quah, D. (1999). *The new empirics of economic growth*. In J. Taylor and M. Woodford (Eds.), *Handbook of macroeconomics*.

Ermisch, J., Gambetta, D., Laurie, H., Siedler, T. and Uhrig, S. (2009). *Measuring people's trust*. *Journal of the Royal Statistical Society*, *172*(4), 749–69.

Fehr, E. (2009). *Economics and biology of trust*. *Journal of the European Economic Association*, *7*(2-3), 235–66.

Fehr, E., Fischbach, U., Rosenbladt, B., Schupp, J. and Wagener, G. (2002). *A nation-wide laboratory: Examining trust and trustworthiness by integrating behavioral experiments into representative surveys*. *Schmollers Jahrbuch*, *122*(4), 519–42.

Felten, E. (2000). *Finders keepers?* *Reader's Digest*, *158*(948).

Fernandez, R. (2011). *Does culture matter?* In J. Benhabib, A. Bisin and M. Jackson (Eds.), *Handbook of social economics, volume 1a.*

Gambetta, D. (2000). *Can we trust trust?* In D. Gambetta (Ed.), *Trust: Making and breaking cooperative relations* (pp. 213–237). Blackwell.

Giavazzi, F., Petkov, I. and Schiantarelli, F. (2019). *Culture: persistence and evolution*. *Journal of Economic Growth*, *24*(2), 117–54.

Glaeser, E., Laibson, D., Scheinkman, J. and Soutter, C. (2000). *Measuring trust*. *Quarterly Journal of Economics*, *115*(3), 811–46.

Glaeser, E., Liabson, D. and Sacerdote, B. (2002). *An economic approach to social capital*. *The Economic Journal*, *112*, 437–58.

Guiso, L., Sapienza, P. and Zingales, L. (2009a). *Cultural biases in economic exchange*. *Quarterly Journal of Economics*, *124*(3), 1095–31.

Guiso, L., Sapienza, P. and Zingales, L. (2009b). *Does Culture Affect Economic Outcomes?* *Journal of Economic Perspectives*, *20*(2), 23–48.

Guiso, L., Sapienza, P. and Zingales, L. (2011). *Civic capital as the missing link*. In J. Benhabib, A. Bisin and M. Jackson (Eds.), *Handbook of social economics, volume 1a.*

Guiso, L., Sapienza, P. and Zingales, L. (2016). *Long-term persistence*. *Journal of the European Economic Association*, *14*(6), 1401–36.

Hardin, R. (2002). *Trust and Trustworthiness*. New York, NY: Russel Sage Foundation.

Herrmann-Pilath, C. (2010). *What have we learned from 20 years of economic research into culture?* *International Journal of Cultural Studies*, *13*(4), 317–35.

Holm, H. and Danielson, A. (2005). *Tropic trust versus Nordic trust: Experimental evidence from Tanzania and Sweden*. *The Economic Journal*, *115*(503), 505–32.

Ikeda, S. (2008). *The meaning of "social capital" as it relates to the market process*. *The Review of Austrian Economics*, *21*(2), 167–82.

Iwashyna, T., Christakis, N. and Becker, L. (1999). *Neighborhoods matter: A population-based study of provision of cardiopulmonary resuscitation*. *Annals of Emergency Medicine*, *4*(1), 459–68.

Jones, G. (2016). *Hive mind: How your nation's IQ matters so much more than your own*. Stanford, CA: Stanford University Press.

Kapas, J. (2017). *How cultural values affect economic growth: A critical assessment of the literature*. *Ekonomska misao i praksa*, *1*, 265–85. Retrieved from https://hrcak.srce.hr/183562

Karlan, D. (2005). *Using experimental economics to measure social capital and predict financial decisions*. *American Economic Review*, *95*(5), 1688–99.

Knack, S. and Keefer, P. (1997). *Does social capital have an economic payoff? A cross-country investigation*. *The Quarterly Journal of Economics*, *112*(4), 1251–88.

Miller, A. and Mitamura, T. (2003). *Are surveys on trust trustworthy?* *Social Psychology Quarterly*, *66*(1), 62–70.

Moore, M. (1999). *Truth, trust and market transactions: What do we know?* *Journal of Development Studies*, *36*(1), 74–88.

Müller, D., Torgler, B. and Uslaner, E. (2012). *A Comment on "Inherited Trust and Growth"*. *Economics Bulletin*, *32*(2), 1481–88. Retrieved from https://ideas.repec.org/a/ebl/ecbull/eb-12-00319.html

Näf, M. and Schupp, J. (2009). *Can we trust the trust game ? a comprehensive examination*. Retrieved from https://www.semanticscholar.org/paper/Can-we-trust-the-trust-game-A-comprehensive-%E2%88%97-N%C3%A4f-Schupp/93902e8d9d80c62db176dce99f1263247e1806de

Nunn, N. (2008). *The long-term effects of Africa's slave trades*. *The Quarterly Journal of Economics*, *123*(1), 139–176.

Nunn, N. and Wantchekon, L. (2011). *The slave trade and the origins of mistrust in Africa*. *American Economic Review*, *101*(7), 3221–52.

Oosterbeek, H., Sloof, R. and van de Kuilen, G. (2004). *Cultural differences in ultimatum game experiments: Evidence from a meta-analysis*. *Experimental Economics*, *7*(2), 171–88.

Putnam, R. D. (2000). *Bowling Alone: The Collapse and Revival of American Community*.

Roodman, D., Nielsen, M. Ø., MacKinnon, J. G. and Webb, M. D. (2019). *Fast and wild: Bootstrap inference in stata using boottest*. *The Stata Journal*, *19*(1), 4–60.

Rose, D. C. (2011). *The Moral Foundation of Economic Behavior*. Oxford, UK: Oxford University Press.

Rose, D. C. (2019). *Why Culture Matters Most*. Oxford, UK: Oxford University Press.

Roth, F. (2009). *Does too much trust hamper growth?* *Kyklos*, *62*(1), 103–28.

Sampson, R. (2012). *Great American City*. Chicago, IL: University of Chicago Press.

Sapienza, P., Toldra-Simats, A. and Zingales, L. (2013). *Understanding trust*. *The Economic Journal*, *128*(873), 1313–32.

Simpson, B., McGrimmon, T. and Irwin, K. (2008). *Are blacks really less trusting than whites? Revisiting the race and trust question*. *Social Forces*, *86*(2), 525–52.

Smith, T. (1997). *Factors relating to misanthropy in contemporary American society*. *Social Science Research*, *26*(2), 170–96.

Smith, T. W., Davern, M., Freese, J. and Morgan, S. L. (2018). *General social surveys, 1972-2018 [machine-readable data file]*. Retrieved from http://gss.norc.org/

Tabellini, G. (2008). *Institutions and culture*. *Journal of the European Economic Association*, *6*(2-3), 255–94.

Tabellini, G. (2010). *Culture and institutions: Economic development in the regions of Europe*. *Journal of the European Economic Association*, *8*(4), 677–716.

Uslaner, E. (2002). *The Moral Foundations of Trust*. Cambridge, UK: Cambridge University Press.

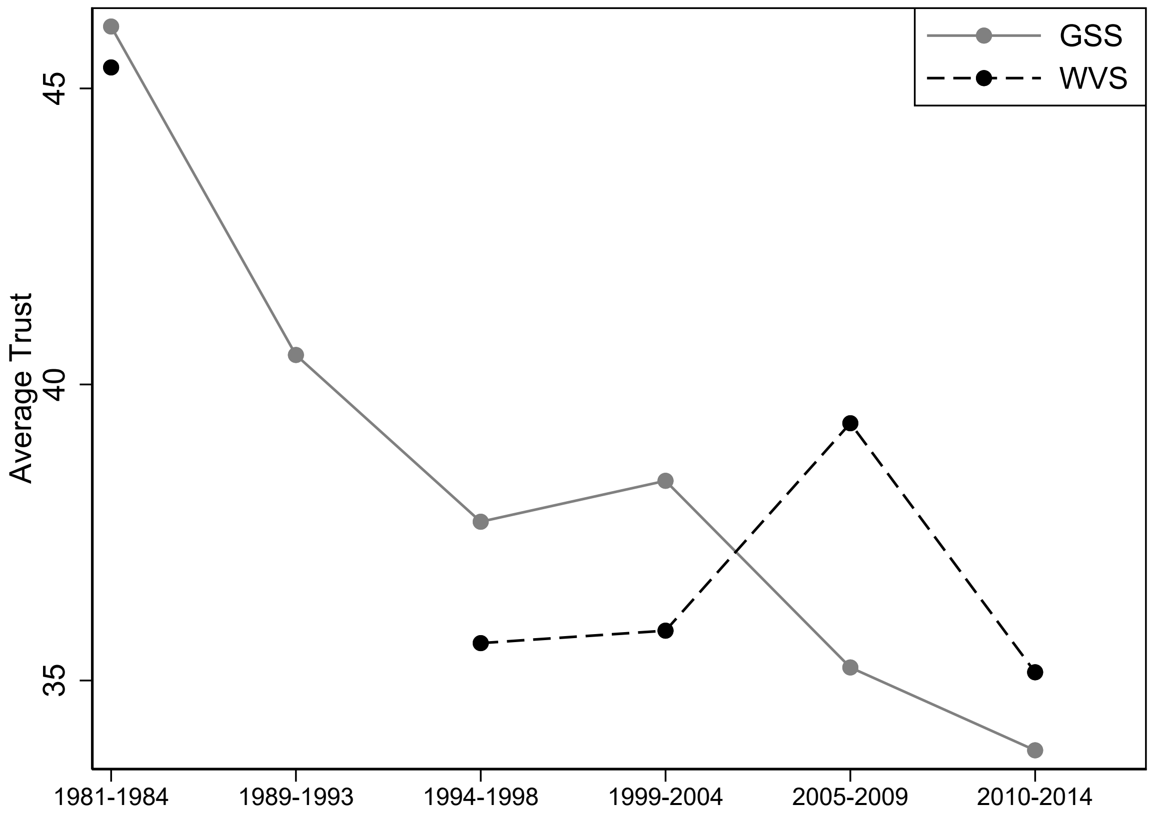
Uslaner, E. (2008). *The foundations of trust: Macro and micro*. *Cambridge Journal of Economics*, *32*(2), 289–94.

Weil, D. (2005). *Economic Growth*. Boston, MA: Addison-Wesley.

Zak, P. and Knack, S. (2001). *Trust and growth*. *The Economic Journal*, *111*(470), 295–321.

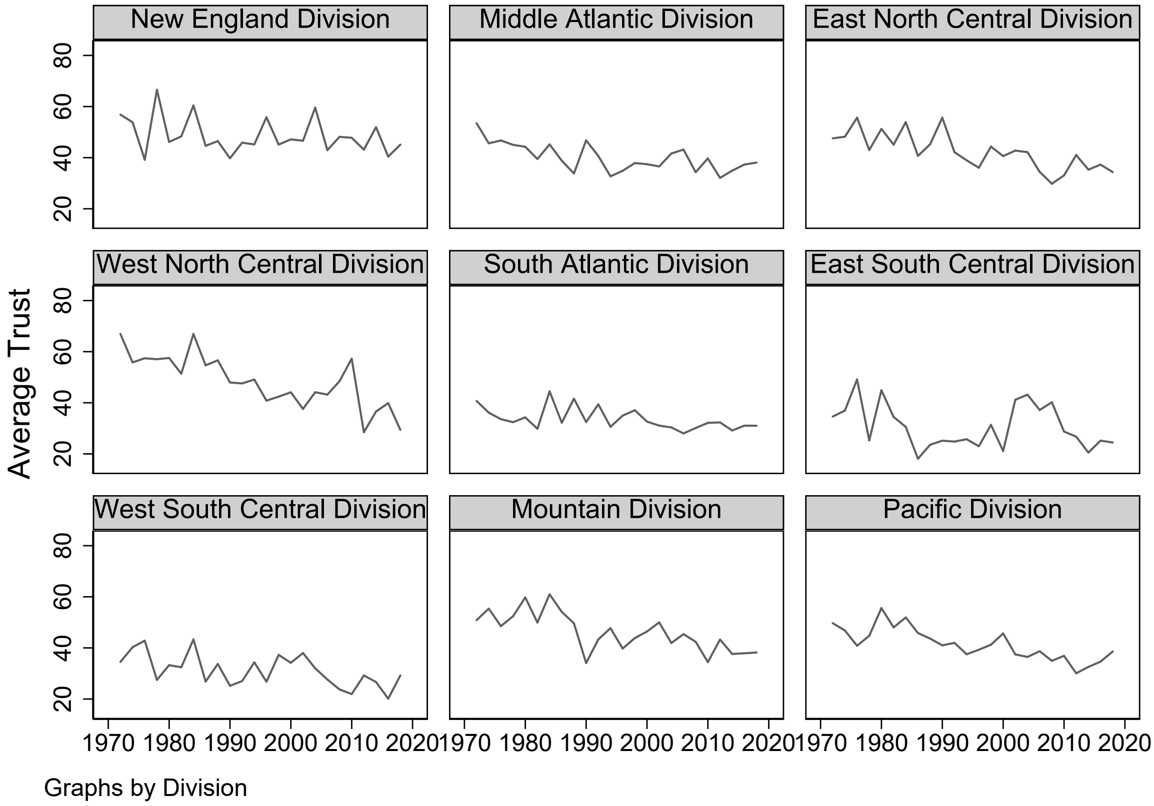
Figures

Figure 1. Average Trust in the U.S., GSS v. WVS



Notes: Figure shows the comparison between average trust between the World Values Survey (WVS) and General Social Survey (GSS). Each measure is computed using survey weights provided in each dataset. The U.S. data were missing for Wave 2 of the WVS.

Figure 2. Average Trust in the U.S., by Census Division



Notes: Figure plots the time series of average trust for each of the 9 Census divisions from the General Social Survey (GSS).

Tables

Table 1. Generalized Trust, by Census Division

	Mean	Std. Dev.	Min	Max
East North Central Division	41.96	6.337	29.82	55.66
East South Central Division	30.55	7.656	18.12	49.14
Middle Atlantic Division	39.88	4.665	32.10	53.49
Mountain Division	46.30	6.817	34.09	60.99
New England Division	48.66	6.595	28.97	66.56
Pacific Division	41.56	6.056	30.10	55.59
South Atlantic Division	33.56	4.117	27.59	44.47
West North Central Division	48.21	9.349	28.49	66.95
West South Central Division	31.08	6.092	20.17	51.12
Total	40.20	9.313	18.12	66.95

Notes: Table presents summary statistics on average generalized trust by Census Division.

Table 2. Trust and Growth, Biennial Data

	(1)	(2)	(3)	(4)
Log Trust ($t-1$)	-0.013	0.004	-0.009	-0.008
	(0.490)	(0.966)	(0.359)	(0.727)
	[-.02193, .01652]	[-.08192, .04606]	[-0.0211, 0.00966]	[-0.0528, 0.0375]
Log Income ($t-1$)			-0.117	-0.123
			(0.548)	(0.000)
			[-0.0294, 0.0246]	[-0.139, -0.106]
Specification	Year FE	Linear Trends	Year FE	Linear Trends
Adj. R-Sq.	0.966	0.653	0.968	0.779
Divisions	9	9	9	9
N	207	207	207	207

Notes: The dependent variable is the log difference of per capita personal income. Log trust denotes the natural log of average generalized trust from the GSS. Log income denotes the natural log of per capita personal income from the BEA. Wild cluster bootstrap p-values are shown below coefficient estimates in parentheses and 95% confidence intervals are shown in brackets.

Table 3. Trust and Growth, Annual Data

	(1)	(2)	(3)	(4)
Log Trust $(t-1)$	-0.004	0.016	-0.002	0.009
	(0.7518)	(0.2823)	(0.587)	(0.221)
	[-.01168, .01084]	[-.01832, .03826]	[-0.0109, 0.00752]	[-0.00864, 0.0326]
Log Income $(t-1)$			-0.046	-0.061
			(0.514)	(0.000)
			[-0.0106, 0.0130]	[-0.0701, -0.0527]
Specification	Year FE	Linear Trends	Year FE	Linear Trends
Adj R-Sq.	0.954	0.582	0.955	0.684
Divisions	9	9	9	9
N	414	414	414	414

Notes: The dependent variable is the log difference of per capita personal income. Log trust denotes the natural log of average generalized trust from the GSS. Log income denotes the natural log of per capita personal income from the BEA. Wild cluster bootstrap p-values are shown below coefficient estimates in parentheses and 95% confidence intervals are shown in brackets.