CATO INSTITUTE

Who Benefits When Firms Game Corrective Policies?

By Mathias Reynaert, Toulouse School of Economics, University of Toulouse Capitole; and James M. Sallee, University of California, Berkeley

Sometimes firms comply with a regulation by gaming the measure targeted by policy rather than by changing their behavior. This relates to Goodhart's Law, which posits that "when a measure becomes a target, it ceases to be a good measure." We study how gaming of this sort can impact consumers who rely on manipulable measures for making choices in a market. Gaming erodes consumer information and induces mistakes. But when the gaming is done in reaction to a regulatory constraint, it lowers the costs to firms, which benefits consumers via pass-through. The impact of gaming on consumer welfare is thus ambiguous, even when the gaming completely fools consumers, and the net effect depends critically on whether gaming is done in response to a policy.

We explore the impact of gaming on consumer welfare both theoretically and empirically for the case of automobile fuel-consumption ratings. We do three things: First, we use a novel data set to measure on-road fuel consumption and document gaming of fuel-consumption ratings, which escalates dramatically following the introduction of regulations that target this rating. Second, we develop a theoretical model that derives the impact of gaming on buyer welfare in a setting where sellers game energy-efficiency ratings, which buyers use to evaluate products. Third, we conduct welfare analysis using estimates of the automobile demand system to quantify the welfare effects identified by our theory.

Our empirical analysis considers the introduction of stringent corrective policy in the EU automobile market. Prior to 2007, there were no policies in Europe that hinged directly on fuel-consumption ratings; since then, both EU standards and nation-specific tax schemes have created policy incentives that reward lower laboratory fuel-consumption test ratings. To measure gaming, we compare the laboratory ratings, which form the basis of policy, with direct measures of on-road fuel consumption that we construct from a data set that tracks fuel consumption and kilometers traveled for a panel of more than 250,000 drivers for a period of 12 years in the Netherlands. Using these data, we estimate the percentage difference between the laboratory test and on-road performance (which we call the performance gap) for each vehicle vintage and model.

We document a sharp rise in the performance gap coincident with policy change. Vehicles produced before 2007 show a small, relatively stable performance gap. Vehicles produced after that exhibit a large and rising performance gap, so that 2014 model-year vehicles have performance gaps in excess of 50 percent on average. The rise in the performance gap implies that around 65 percent of the gains in fuel economy since the introduction of policy, as measured by laboratory tests, are false. Using conventional estimates of lifetime distance traveled and a social cost of carbon at $40 per ton, the difference between apparent and actual emission reductions amounts to $1.2 billion annually from 2010 to 2014 when extrapolated to

all of Europe. We interpret the rise in the performance gap as evidence of gaming in response to policy incentives, in the spirit of Goodhart's Law, and then turn our attention both theoretically and in welfare simulations to our central question: Who benefits when firms game corrective policy?

Our theoretical model considers a monopolist who sells a good to a representative consumer. The good has some attribute that is desirable to consumers, but it also creates a negative externality that motivates corrective policy. The attribute closely matches the role of fuel economy ratings in the automobile market. The attribute is not directly observable, however, so consumer demand and government regulation are based on a measure provided by the seller. The seller can change the measure either by changing the true attribute or by gaming, both of which are costly.

In our model, we allow that some fraction of gaming is undetected by buyers. In the absence of policy, this means that gaming lowers buyer welfare for two reasons. First, gaming causes buyers to misoptimize (choose the wrong quantity of the good), which leads to a loss in buyer surplus that we call choice distortion. Second, gaming causes the seller to raise prices because buyers perceive an improvement in the product. This price effect further reduces buyer surplus.

Corrective policy disrupts this logic by flipping the sign of price effects. Regulation raises the cost of production. Gaming allows the seller to lower its costs, and this benefits buyers through lower prices in the same way a reduction in a tax would. When this price effect dominates choice distortions from faulty information, buyers benefit from the seller's gaming even when they are fooled by it.

We focus on buyer surplus as a notion of the private surplus of consumers that consider buying the good. This is narrower than consumer surplus, which would encompass the externality, but gaming will also impact the level of the externality. The ultimate effect of gaming on the environment depends on the sophistication of the policymaker, who may increasingly tighten policy to achieve real gains, and of the buyers, who may expand the overall size of the market when they mistakenly perceive lower costs of ownership.

Next, we set out to quantify the price effect and choice distortions in our empirical setting. We demonstrate that the price effects and choice distortions identified by our theory have direct empirical analogs in a discrete choice setting. We then estimate a demand model of the European car market that provides us estimates of consumer preferences and the marginal costs of products. Given these preferences and costs, we calibrate the incidence of gaming for a range of alternative assumptions regarding consumer awareness, policy stringency, and the degree of gaming.

We find robust results that align with our theoretical predictions. When there is no corrective policy and when consumers are fooled by gaming, we find that lowering perceived fuel costs through gaming leads to modest losses in buyer surplus. A significant majority of these losses come through the price effect. As a result, firm profits rise with gaming, and this comes at the expense of lower buyer surplus.

As suggested by the theory, the welfare effects of gaming change when we introduce a corrective policy. We model a mandated decrease in average fuel-consumption ratings that firms comply with via shifting their sales mix toward more efficient models. When firms comply honestly, private consumer surplus falls substantially, as consumers are forced into less desirable products. When we allow firms to relax this regulatory constraint by gaming to meet the standard, we find beneficial price effects for buyers that consistently dominate choice distortions so that the net impact of gaming is to raise buyer surplus. Gaming with or without a policy induces a similarly sized choice distortion, but we find that this is an order of magnitude smaller than the price effect for a significant range of parameter choices. These results provide empirical validation for our theoretical prediction: gaming benefits consumers in the presence of stringent policy, even when buyers are fooled. We also show that a corrective policy roughly triples the private benefit to a single firm that games when all others are honest, which implies that policy amplifies competitive pressures that incentivize gaming.

NOTE:

CATO
INSTITUTE