HYSTERESIS AND THE SOCIAL COST OF CORRECTIVE POLICIES: EVIDENCE FROM A TEMPORARY ENERGY SAVING PROGRAM

Francisco Costa (FGV/EPGE)   François Gerard (Columbia University)

June 4, 2015

Abstract

This paper studies how one may overestimate the social cost of a long-run corrective policy by neglecting the possibility of hysteresis, i.e., that the policy in earlier periods may have a persistent impact in the long run. In a price-theoretic framework, we show that one statistic is key to evaluating this bias: the long-term impact of a similar but temporary policy that was known to be temporary. We then provide evidence of the importance of hysteresis, and estimate such a statistic, for a policy-relevant behavior: residential electricity use in a developing country context. We study the 10-year impact of a 9-month long policy in Brazil, which aimed at large temporary reductions in residential electricity use. Through a difference-in-difference strategy, we exploit the fact that customers of some distribution utilities were not subject to the policy. Using utility-level administrative data, we find that the temporary policy led to a long-run and stable reduction in average electricity use of 11%, or about half of the short-run impact. Using individual monthly billing data for one distribution utility, we find that 69% of customers were still consuming less electricity four years after the policy ended. Household-level microdata suggest that the main mechanism of hysteresis is a persistent change in consumption habits. Incorporating our estimates into this framework illustrates that, by neglecting the possibility of hysteresis, one could dramatically overestimate the social cost of long-run corrective policies. (*JEL:* D62, H23, Q50)

We would like to thank Michael Best, Gharad Bryan, Robin Burgess, Lucas David, Marina de Mello, Ben Faber, Claudio Ferraz, Fred Finan, Greg Fischer, Meredith Fowlie, Jason Garred, Gustavo Gonzaga, Ryan Kellog, Gilat Levy, Jamie McCasland, Guy Michaels, Edward Miguel, Mushfiq Mobarak, Marcelo Moreira, Gerard Padró i Miquel, Emmanuel Saez, Edson Severnini, Reinaldo Souza, Anna Spurlock, Eric Verhoogen, Catherine Wolfram, and seminar participants at ASU, the Behavioural Environmental Economics conference (Toulouse), Cergy-Pontoise, EconCon, the EEA meetings, FGV/EPGE, FGV/EESP, the Granthan Institute for Climate Change, IADB, IFS, IPEA, LBS, LSE, the NBER Public Economics Meetings, PacDev, PSE, Puc-Rio, UC Berkeley, University of Bristol, University of Cambridge, University of Chicago, UPF, UQAM, USP, UT Knoxville, and the World Bank. Stéphanie Dinóa, Tiago Lazier, Gustavo Macedo, and John Pease provided outstanding research assistance. We also thank ANEEL, PROCEL, LIGHT, FAME, and Whirlpool for sharing important data. François Gerard benefited from the support of an Excellence Scholarship of Wallonie-Bruxelles International. Francisco Costa, Getulio Vargas Foundation, FGV/EPGE - Escola Brasileira de Economia e Finanças, e-mail: francisco.costa@fgv.br. François Gerard, Columbia University, Department of Economics, e-mail: fgerard@columbia.edu.

Policymakers may want to induce long-run changes in the behavior of economic agents in the presence of externalities, e.g., reduce energy demand to mitigate environmental damages. However, any corrective policy will create a social cost that depends on the size of the change in behavior and on agents' costs of changing their behavior. The latter is generally captured by the price elasticity of demand (or supply) for the behavior. Therefore, inducing large long-run changes in a behavior that presents a low elasticity may imply a high social cost in every period, and may not be desirable despite sizable externalities. This influential argument rests on the assumption that the level of a behavior depends only on the concurrent economic environment. It leaves out the possibility that the impact of the corrective policy in earlier periods may change the level of a behavior in a way that persists in the long run even in the absence of the policy. This paper studies the importance of such hysteresis for the social cost of long-run corrective policies.

Hysteresis is defined as the failure of an effect to reverse itself as its underlying cause is reversed (Dixit, 1989). We use the term hysteresis to refer to the possibility that the impact of a temporary policy could persist in the long run or that there may be multiple steady states for a given behavior. We emphasize *long-run persistence* because observing a delay before a behavior returns to its prior level once a policy is suspended could simply result from a difference between short- and long-run elasticities. The implications are well-known in this case: long-run elasticities should be used to evaluate the social cost of long-run policies. Several theories allow for hysteresis,[1] but there is limited evidence that it is an empirically relevant phenomenon in many contexts of interest.

We proceed in steps. First, we use a price-theoretic framework to illustrate how one could overestimate the social cost of a long-run corrective policy by neglecting the possibility of hysteresis. Quantifying the degree of hysteresis is key to evaluating this bias. We argue that this can be done by estimating the long-term impact of a similar but temporary policy that agents expected to be temporary. The core of the paper then provides evidence of the importance of hysteresis, and estimates such a statistic, in a policy-relevant behavior: residential electricity use in a developing country context. We estimate the 10-year impact of a 9-month long electricity saving program that was imposed on millions of Brazilian households in 2001 and that led to dramatic short-run reductions in residential electricity use. Exploiting administrative data and differences in the implementation of the policy across regions, we find that about half of the short-run impact persisted in the long run. Household level microdata suggest that the main mechanism of hysteresis was a change in electricity utilization habits. Finally, we incorporate our estimates into this conceptual framework to illustrate implications for the social cost of corrective policies.

We begin by using a simple partial-equilibrium framework to show why hysteresis matters for

[1] Any model with asymmetric adjustment costs can generate hysteresis. For example, by having to change their behavior, agents may learn new ways to derive more utility from a given level of behavior. Habit formation à la Becker and Murphy (1988) allows for, but does not imply, the existence of multiple steady states. Therefore, it certainly implies a difference between short- and long-run elasticities, but may or may not imply hysteresis.

the social cost of corrective policies. The intuition is straightforward. Consider a policymaker that imposes a permanent change in behavior, e.g., a 20% reduction in residential electricity use. In the absence of hysteresis, one would assume that the policy distorts behavior by 20% in all periods. Yet, with hysteresis, part of the reduction in later periods would be caused by the policy in earlier periods and should not be double-counted as caused by the concurrent policy. An estimate of the *long-term* impact of a similar but temporary policy that agents *expected to be temporary* quantifies the degree of hysteresis and thus the residual distortion that the permanent policy would cause in later periods. Studying the long-term impact rules out any persistence resulting from a difference between short- and long-run price elasticities. Studying a policy that was known to be temporary rules out any persistence that would not be due to the temporary policy, but to past responses in anticipation of a continuation of the policy. Integrating the demand (or supply) curve over the estimated residual distortion vs. the 20% reduction evaluates how one would overestimate the social cost of the policy in later periods by neglecting the possibility of hysteresis.

We then study the long-term impact on residential electricity use of the temporary electricity saving program that was implemented during the 2001 Brazilian electricity crisis. Our empirical setting is ideal to provide evidence of hysteresis because the policy was known to be temporary and because we can study impacts up to ten years after the policy ended. The Brazilian program also led to the largest short-run reductions in household electricity use among temporary electricity saving programs around the world (Meier, 2005). This may be important because several theories suggest that hysteresis is more likely to be relevant for large changes in behavior.[2]

Our empirical setting is also relevant in itself. First, energy use is expected to continue to be a major source of greenhouse gas emissions in the future. Potential energy savings from residential electricity use have attracted a lot of attention.[3] Yet, existing estimates of price elasticities are typically low. Thus, inducing large long-run changes in residential electricity use is an example where the social cost may be considered too large despite sizable externalities. Second, most of the growth in energy use is forecast to come from the developing world. In particular, greenhouse gas emissions from residential electricity use are growing rapidly (IPCC, 2014). Yet, the energy saving potential of households in developing countries, who are poorer, own fewer appliances, are more credit-constrained, and consume less energy to begin with, is largely unknown.

In the beginning of 2001, electricity generation capacity was severely reduced in some regions

[2]For instance, learning effects are more likely for policies that force agents to depart further from their usual levels of behavior. There may also be important fixed costs. In a rational habit formation model, hysteresis relies on pushing agents far enough from their prior steady state (Becker and Murphy, 1988). Other temporary electricity saving programs led to smaller reductions in average electricity use at a given point in time and/or were shorter-lived.

[3]Improving the energy efficiency of residential electricity use is often viewed as a cost-effective policy to abate greenhouse gas emissions (McKinsey, 2009). Utilities have to meet specific energy saving targets through customer electricity saving programs in at least 24 states in the US. Scenarios to mitigate the impacts of climate change typically involve large reductions in electricity use by buildings (IPCC, 2014).

of Brazil. A temporary shock to the streamflow level in the rivers that serve the hydroelectric power plants in these regions led to historically low water levels in their reservoirs (see Figure 1a). Brazil relies heavily on hydroelectric generation and low transmission capacity constrained transfers of electricity across regions. In order to prevent generalized blackouts, the government implemented a temporary electricity saving program from June 2001 to February 2002 in affected areas, which aimed at reducing residential electricity use by 20%. Residential customers were assigned individual quotas and were subject to a series of incentives to consume below their quota.[4]

We estimate the short- and long-term impacts through a difference-in-difference approach, comparing distribution utilities subject to the policy to those that were exempt. We use data on average residential electricity use per customer from monthly administrative reports for every distribution utility between 1991 and 2011. We confirm that the policy had a large short-run impact (-23%). Our main contribution is to show that half of that impact persisted in the long run (-11%).[5] Consumption levels partially rebounded after the policy ended but point estimates are stable from 2005 onward; this can be seen in the raw data in Figure 1b. It is thus unlikely that our estimates are due to a difference between short- and long-run elasticities rather than to hysteresis.

We present many empirical tests supporting our results. Our estimates are robust to controlling for a series of confounders such as changes in electricity tariffs, demographics (Levinson, 2014), income levels, or climate. Moreover, utility-specific impacts estimated by synthetic control methods find negative long-term impacts for *every* distribution utility subject to the policy. We also rely on longitudinal monthly billing data from 2000 to 2005 for three million customers of one affected utility. We show that changes in average electricity use are similar in the aggregate data and in a random sample of individual customers observed every month in the billing data. By construction, this balanced panel is free of composition effects. Almost all of these customers consumed less electricity during the crisis than they did a year before (92%) and most of them continued to consume less electricity four years after the crisis (69%). The median customer reduced electricity use by 31% and 16.5% during the crisis and four years after, respectively. These figures are larger for customers with higher baseline consumption and there is a strong correlation at the customer level between short-run (during the crisis) and long-run (four years later) changes in consumption. Average effects thus came from large and persistent responses by most customers. Finally, household

[4]Incentives cannot be translated into a given increase in tariffs because they included nonlinear pecuniary incentives (fines or bonuses for consuming above or below the quota) and non-pecuniary incentives (threats of disconnection for consuming above the quota and moral suasion from conservation appeal campaigns). We study the overall impact of the policy because we cannot separately estimate the impacts of its components.

[5]There is little work on the long-run impact on residential electricity use of the 2001 Brazilian electricity crisis and the associated electricity saving program. Bardelin (2004) and Maurer, Pereira and Rosenblatt (2005) provide some descriptive evidence on short-run impacts with aggregate data. Pimenta, Notini and Maciel (2009) use time-series techniques. Mation and Ferraz (2011) use a similar difference-in-difference strategy to investigate impacts on firms' productivity. In a different context, Reiss and White (2008) present time-series evidence on how households responded to price increases and conservation appeals during the 2000 California energy crisis.

surveys of appliance ownership and utilization habits conducted both before the crisis and several years later suggest that the main mechanism of hysteresis was the formation of new habits.

We recognize that there may be concern about the external validity of our evidence. This is an issue in any empirical setting (Allcott, forthcoming), which may be seen as more severe in our case given the uniquely large short-term changes in behaviors. However, the uniqueness of our setting does not come from a lack of interest from policy and academic circles in policies aimed at large changes in behaviors. It comes from the fact that such policies are rarely implemented. It is therefore particularly interesting to exploit an opportunity to learn from such a policy. This is particularly true if the reason that these policies are considered politically infeasible is because their social cost is assumed to be too large. Furthermore, the fact that we find persistent effects for all affected utilities, which differ in the characteristics of their local demand, and across the distribution of the millions of customers of one utility, brings some external validity to our evidence.

This paper contributes to a large literature on the economics of corrective policies, dating back at least to Pigou (1920). The existing literature typically fails to consider the role of hysteresis for policies that aim at changing behaviors in the long run.[6] In so doing, we show that one could largely overestimate the social cost of long-run corrective policies and we identify estimable "sufficient statistics" to evaluate this bias.[7] Moreover, our analysis of the Brazilian program indicates that taking hysteresis into account can be quantitatively important in a policy-relevant context. A back-of-the-envelope calculation implies that, neglecting the possibility of hysteresis, one could overestimate the social cost of a 10-year long version of a similar policy by 182%.

This paper also contributes to the growing empirical literature that investigates the presence and mechanisms of persistence in various behaviors of interest. Several papers study the persistence of some impacts after a policy was suspended.[8] It is sometimes unclear whether their estimates actually capture hysteresis. Agents may have expected the policy to continue with some probability and the persistence is often observed for a relatively short period of time.[9] Moreover, to the extent that some do capture hysteresis, one could use our framework to evaluate the overestimation bias from neglecting hysteresis in calculating the social cost of long-run corrective policies in their

[6]In recent theoretical work, Acemoglu et al. (2012) argue that temporary policies promoting greener technologies may have persistent effects on the supply side through directed technical change and Hintermann and Lange (2013) investigate the dynamic optimal regulation of experience goods generating environmental externalities.

[7]The same argument would apply to any corrective policy that aims at changing two complementary behaviors; hysteresis essentially implies that behaviors in different periods are complementary. We are not aware of existing work applying this idea to intertemporal behaviors.

[8]E.g. Charness and Gneezy (2009), Giné, Karlan and Zinman (2010), Ferraro and Price (2013), Bryan, Chowdhury and Mobarak (2014), Dupas (2014), Fujiwara, Meng and Vogl (2014), Acland and Levy (2015), Miller (2014), Allcott and Rogers (2014) and Ito, Ida and Tanaka (2015).

[9]For instance, Allcott and Rogers (2014) find a persistent change in electricity use after some incentives were suspended and estimate that it would take about five years for the effect to disappear in their context. They had access to only two years of post-intervention data, however. We would have reached a similar conclusion with a similar data limitation. Instead, we can show that average electricity use does not return to counterfactual levels in our context.

context. Existing studies do not consider this implication of their results.

Finally, this paper contributes to a small but growing empirical literature that studies issues related to energy consumption in developing countries.[10] We show that a temporary shock that forced Brazilian households to conserve electricity impacted consumption levels in the long run. Moreover, this hysteresis seems to arise mostly from the formation of new habits rather than from physical investments.[11] Our results thus open up an exciting research agenda: how could policy-makers foster energy-efficient habits at an earlier stage of development to limit the rapidly growing energy demand in the developing world.

The rest of the paper proceeds as follows. Section 1 presents our conceptual framework. Section 2 introduces our empirical setting and Section 3 our empirical strategy. Section 4 presents our main results and the many robustness checks supporting them. Section 5 investigates the mechanisms of hysteresis. Section 6 uses our estimates to illustrate the importance of taking hysteresis into account when estimating the social cost of long-run corrective policies. Section 7 concludes.

1 Conceptual framework

In this section, we provide a theoretical framework to illustrate how one may overestimate the social cost of long-run corrective policies by neglecting the possibility of hysteresis in the behavior of interest.[12] We present here the simplest version of the model. We discuss some extensions, but leave the related derivations to the Web Appendix.

Hysteresis could emerge in several models, which would predict the same observable outcome. It could emerge in models with asymmetric adjustment costs if agents pay some costs that cannot be recovered (Dixit, 1989). For example, a corrective policy may cause agents to learn about the costs and benefits of adopting new behaviors (Bryan, Chowdhury and Mobarak, 2014; Dupas, 2014) or to learn new ways to derive more utility from a given level of behavior. Hysteresis could also occur in a rational habit formation model à la Becker and Murphy (1988) if there are multiple

[10] Wolfram, Shelef and Gertler (2012) argue that existing forecasts underestimate the future growth in residential energy demand in the developing world. They also highlight that rapidly rising energy demand brings the risk of dramatic supply shortages in developing countries because of vulnerable infrastructure and the difficulty of accurately planning capacity investments. Davis, Fuchs and Gertler (2014) find that an appliance replacement program in Mexico did not generate the expected energy saving because of households' behavioral responses to the new appliances. Allcott, Collard-Wexler and O'Connell (2014) and Fisher-Vanden, Mansur and Wang (2015) examine firms' short-run responses to recurrent power shortages in India and China, respectively.

[11] This should not be surprising. Households in developing countries are poorer and more credit-constrained, and habit formation is relevant for residential electricity use even in developed countries (e.g., Ito, Ida and Tanaka, 2015).

[12] We could instead consider the role of hysteresis for the welfare effects of temporary policies. In the presence of hysteresis, a temporary policy will deliver a persistent correction of the externality, if the prior level of the behavior was too high from a social perspective. Properly accounted for, this may partly mitigate the social cost borne by economic agents in moving between steady states. This does not imply, however, that a temporary policy would be desirable, as the level of the behavior in the new steady state may still be suboptimal from a social perspective.

steady states and a policy pushes agents far enough from their prior steady state. We do not take a stand on the underlying model at play. We simply assume that policies that affect past levels of a behavior may affect agents' utility from different levels of the behavior in the long run.[13] We then illustrate the importance of hysteresis assuming that price theory can be used to evaluate the social cost of corrective policies. Alternatively, hysteresis could emerge from "behavioral" theories, such as models with biased beliefs or myopia about the returns to changing one's behavior (e.g., Gruber and Köszegi, 2001; Acland and Levy, 2015). We return to this point later in this section.

1.1 Setup

We adopt a series of assumptions to focus on the role of hysteresis. We consider a representative-household two-period model of electricity use.[14] The periods are assumed to be long enough that the level of electricity use in the two periods would be independent in the absence of hysteresis, i.e., there is no difference between short- and long-run elasticities. We adopt a partial-equilibrium setup, we assume away income effects and redistributive concerns, and we model the household as fully rational and forward looking. We also assume that goods are produced competitively at constant marginal costs. These assumptions allow us to illustrate the role of hysteresis for the social cost of long-run corrective policies using simple consumer surplus concepts.

The household chooses electricity consumption x_i at price p_i and ordinary consumption c_i at normalized price 1, given income y_i in periods $i = 1,2$. The per-period utility is represented by $u_i(c_i, x_i, s_i) = c_i + v_i(x_i, s_i)$, where s_i is the household's propensity to consume electricity. This is a reduced-form variable that is aimed at capturing possible mechanisms of hysteresis. Changes in electricity use may affect the future propensity to consume because the household develops persistent new consumption habits, learns about ways to consume electricity more efficiently, develops a taste for appliances with different characteristics, etc.[15]

We assume that $v_i(x_i, s_i)$ is strictly increasing and concave in x_i and that $\frac{\partial v_i(x_i,s_i)}{\partial s_i} \leq 0$ for all $x_i, s_i \geq 0$. For instance, the household derives less utility from a consumption bundle (c_i, x_i) if her habits, or lack of knowledge about energy-efficient behaviors, require more electricity to provide

[13] Our approach to evaluating the deadweight loss of a corrective policy applies to a wide range of models because it relies on price theory. However, depending on the specific model at play, alternative policies may achieve an equal change in behavior at a lower social cost. This would be the case, for instance, if agents change their behavior persistently by simply receiving some information (without having to experiment with costly changes in their behavior in the first place) that could be provided at low cost. This is unlikely to be the case in our empirical application.

[14] The model can be easily adapted to apply to other behaviors of interest or extended to a multi-period setup.

[15] We allow for direct (costly) investments in one's propensity to consume electricity in an extension of our model (in the Web Appendix). Our approach thus applies to investments in physical capital. However, it is unclear whether investments in physical capital such as energy-efficient appliances would lead directly to hysteresis. A corrective policy may simply push agents to anticipate future investments in new appliances and physical capital decays. Of course, hysteresis may arise if agents' preferences or information sets change when they buy energy-efficient appliances, such that they continue to acquire more efficient appliances subsequently.

the same level of services. The propensity to consume in the first period is given, but, in the second period, it is allowed to depend on the level of electricity use in the first period: $s_2 = s(s_1, x_1)$, with $\frac{\partial s(s_1,x_1)}{\partial s_1}, \frac{\partial s(s_1,x_1)}{\partial x_1} \geq 0$. Finally, we assume that $\frac{\partial^2 v_i(x_i,s_i)}{\partial x_i \partial s_i} > 0$ for all $x_i, s_i \geq 0$: the higher the propensity to consume (e.g., to use electrical appliances), the higher the marginal utility of consumption (e.g., the greater the disutility from not having access to their services). This complementarity introduces some path dependency to the utility derived from electricity use and allows for hysteresis in electricity use. Specifically, the household solves:

$$\max_{c_1,c_2,x_1,x_2} V = u_1 + \beta u_2 = c_1 + v_1(x_1, s_1) + \beta\left[c_2 + v_2(x_2, s(s_1, x_1))\right] \quad \text{s.t. } c_i + p_i x_i \leq y_i$$

where V is the household "lifetime" utility and β accounts for discounting or differences in the relative length of the two periods. Substituting in for c_i, we obtain the first-order conditions:

$$x_1 : \frac{\partial v_1(x_1, s_1)}{\partial x_1} + \beta \frac{\partial v_2(x_2, s(s_1, x_1))}{\partial s} \frac{\partial s(s_1, x_1)}{\partial x_1} = p_1 \quad ; \quad x_2 : \frac{\partial v_2(x_2, s(s_1, x_1))}{\partial x_2} = p_2 \qquad (1)$$

The left- and right-hand sides of each first-order condition capture the benefit and the cost of marginal changes in electricity use, respectively. These must be equal at an optimum. As expected, the household will use less electricity in the first period if the propensity to consume electricity in the second period depends on past choices and the household is aware of this relationship.

1.2 The social cost of a long-run corrective policy

Without government intervention, the first-order conditions and baseline electricity prices in the two periods, p_{10} and p_{20}, will determine baseline electricity use, x_{10} and x_{20}. Now, suppose that the government wants the household to reduce electricity use to $\overline{x_1} < x_{10}$ and $\overline{x_2} < x_{20}$. We are interested in the social cost or deadweight loss of such a corrective policy. In our framework, this is the change in the household's lifetime utility: $DWL = V(\overline{x_1}, \overline{x_2}) - V(x_{10}, x_{20})$. We can recover it by tracing the change in utility along any path from (x_{10}, x_{20}) to $(\overline{x_1}, \overline{x_2})$. In our setting, it is natural to sequentially change x_1 to $\overline{x_1}$ and then x_2 to $\overline{x_2}$, holding constant $\overline{x_1}$. In the presence of hysteresis, we have to take into account the fact that the optimal level of x_2 at price p_{20} will change in the first step, as can be seen from the first-order condition for x_2. We have:

$$
\begin{aligned}
DWL &= \int_{x_{10}}^{\overline{x_1}} \frac{dV(x_1,x_2)}{dx_1} dx_1 + \int_{x_2(\overline{x_1})}^{\overline{x_2}} \frac{dV(\overline{x_1},x_2)}{dx_2} dx_2 \\
&= \int_{x_{10}}^{\overline{x_1}} \left[\frac{\partial v_1(x_1,s_1)}{\partial x_1} + \beta \frac{\partial v_2(x_2,s(s_1,x_1))}{\partial s} \frac{\partial s(s_1,x_1)}{\partial x_1} - p_{10} + \beta \underbrace{\frac{\partial x_2}{\partial x_1} \left[\frac{\partial v_2(x_2,s(s_1,x_1))}{\partial x_2} - p_{20} \right]}_{=0} \right] dx_1 \\
&\quad + \beta \int_{x_2(\overline{x_1})}^{\overline{x_2}} \left[\frac{\partial v_2(x_2,s(s_1,\overline{x_1}))}{\partial x_2} - p_{20} \right] dx_2 \\
&= \int_{x_{10}}^{\overline{x_1}} [p_1(x_1) - p_{10}] \, dx_1 + \beta \int_{x_2(\overline{x_1})}^{\overline{x_2}} [p_2(x_2|\overline{x_1}) - p_{20}] \, dx_2 \qquad (2)
\end{aligned}
$$

where $x_2(\overline{x_1})$ is defined by $\frac{\partial v_2(x_2,s(s_1,\overline{x_1}))}{\partial x_2} = p_{20}$. Equation (2) shows that three empirical objects are sufficient to evaluate the deadweight loss of the policy (besides the parameter β). The first one is the inverse demand curve in period 1, $p_1(x_1)$, which will factor in any effect on utility and behavior in period 2 from changes in x_1. It can be identified using exogenous price increases in period 1. The second one, $p_2(x_2|\overline{x_1})$, is the inverse demand curve in period 2 for a given value of $\overline{x_1}$, and thus a given propensity to consume in period 2. It can be identified using exogenous price increases once in period 2, following a temporary policy that changes x_1 to $\overline{x_1}$. The third one, $x_2(\overline{x_1})$, is the degree of hysteresis, the change in x_2 caused by the change in x_1. It could be identified from the impact on electricity use in period 2 of a similar temporary policy. Equation (2) is a familiar expression for the social cost of a behavioral change. It corresponds to the sum of the area under the properly defined inverse demand curve and above the baseline price level in each period. In the second period, the integral is only taken over the residual change in quantity because any change in utility from changing x_2 to $x_2(\overline{x_1})$ is already accounted for in the first integral. The same argument would apply to any two behaviors that are complementary: hysteresis essentially implies that behaviors at different points in time are complements.[16]

Assuming away hysteresis ($\frac{\partial s(s_1,x_1)}{\partial x_1} = 0$), the expression for the deadweight loss would be:

$$
\begin{aligned}
DWL_{NoH} &= \int_{x_{10}}^{\overline{x_1}} \left[\frac{\partial v_1(x_1,s_1)}{\partial x_1} - p_{10} \right] dx_1 + \beta \int_{x_{20}}^{\overline{x_2}} \left[\frac{\partial v_2(x_2,s_1)}{\partial x_2} - p_{20} \right] dx_2 \\
&= \int_{x_{10}}^{\overline{x_1}} [p_{1,NoH}(x_1) - p_{10}] \, dx_1 + \beta \int_{x_{20}}^{\overline{x_2}} [p_{2,NoH}(x_2) - p_{20}] \, dx_2 \qquad (3)
\end{aligned}
$$

[16] Suppose, for instance, that one smokes only when drinking. A policy that reduces drinking would thus reduce smoking. The demand curve for drinking captures the associated utility from smoking and so this demand curve is sufficient to measure any change in utility from the associated change in smoking. Now suppose that a policy also reduces smoking below the initial level. The reduction in smoking that took place because of the reduction in drinking should not be double-counted. Similarly, if their complementarity were neglected, one would underestimate the change in the levels of both behaviors following an increase in the price/cost of both behaviors (e.g., corrective taxes).

where *NoH* stands for "no hysteresis". There are two sources of bias when assuming away hysteresis. First, one may use the wrong shapes for the demand curves by assuming that demand curves are independent in each period. It is unclear in which direction the bias would go. Assuming away hysteresis, one would identify the inverse demand curve $p_{1,NoH}(x_1)$, using the same variation that identifies the inverse demand curve $p_1(x_1)$. Moreover, it is unclear whether the price elasticity would increase or decrease in the second period, following a temporary corrective policy that leads to a persistent impact.[17] Second, one would neglect the fact that the optimal level of x_2 will change with the level of x_1. This will bias the deadweight loss upward (in absolute values) as it implies taking the integral in the second period over a larger interval. In this paper, we focus on the second source of bias. This is because there is limited evidence to begin with that hysteresis is an empirically relevant phenomenon in many contexts of interest.[18]

Figure 2 provides a graphical illustration assuming linear demand curves and no bias from having the wrong slopes for the demand curves. Reducing quantity in period 1 to $\overline{x_1}$ increases the household's marginal (lifetime) utility for x_1, which can be traced along the inverse demand curve $p_1(x_1)$. In the presence of hysteresis, this would reduce x_2 endogenously to $x_2(\overline{x_1})$. The demand curve $p_1(x_1)$ would factor in any change in utility from such endogenous changes in x_2. The loss in utility from reducing quantity to $\overline{x_1}$ is then the triangle $A_1B_1C_1$. Further reducing x_2 to $\overline{x_2}$ (for a given $\overline{x_1}$) then increases the household's marginal utility for x_2, which can be traced along the inverse demand curve $p_2(x_2|\overline{x_1})$. The associated loss in utility is the triangle $A_2D_2E_2$. Neglecting the possibility of hysteresis, one would overestimate the deadweight loss in the second period. Tracing the change in marginal utility along the whole interval from x_{20} to $\overline{x_2}$, one would obtain a loss in utility corresponding to the larger triangle $A_2B_2C_2$. The bias is the area $D_2B_2C_2E_2$. Formally, assuming linear demand curves and an equal elasticity in the two periods, the bias is:

$$\text{In levels: } |DWL_{NoH}| - |DWL| = \frac{1}{2}\beta\frac{p_{20}x_{20}}{|\eta|}\left[[D]^2 - \left[D - \frac{x_2(\overline{x_1}) - x_{20}}{x_{20}}\right]^2 \frac{1}{x_2(\overline{x_1})/x_{20}}\right] \quad (4)$$

$$\text{In percents: } \frac{|DWL_{NoH}| - |DWL|}{|DWL|} = \frac{\beta p_{20}x_{20}\left[[D]^2 - \left[D - \frac{x_2(\overline{x_1})-x_{20}}{x_{20}}\right]^2 \frac{1}{x_2(\overline{x_1})/x_{20}}\right]}{p_{10}x_{10}[D]^2 + \beta p_{20}x_{20}\left[D - \frac{x_2(\overline{x_1})-x_{20}}{x_{20}}\right]^2 \frac{1}{x_2(\overline{x_1})/x_{20}}} \quad (5)$$

where $D = \frac{\overline{x_1}-x_{10}}{x_{10}} = \frac{\overline{x_2}-x_{20}}{x_{20}}$ is a given long-run change in electricity use (policy goal) and η is the

[17] The elasticity may increase if households start to pay more attention to their electricity use (Jessoe and Rapson, 2014). In this case, assuming away hysteresis would further overestimate the deadweight loss (in absolute values).

[18] An extensive literature focuses on estimating price elasticities for residential electricity use. Long-run price elasticities are always difficult to estimate. In the Web Appendix, we estimate a medium-run price elasticity for the years following the Brazilian temporary policy. We use it when evaluating equation (4) below. Unfortunately, it is not possible to estimate a comparable elasticity for the years leading to (or for the months of) the temporary policy. We are thus unable to document any change in the price elasticity.

price elasticity. Evaluating the bias critically relies on estimating the impact in period 2 of a similar policy implemented only in period 1: $x_2(\overline{x_1}) - x_{20}$. We estimate such a statistic in the remainder of the paper. We then illustrate the importance of hysteresis for the social cost of long-run corrective policies by evaluating equations (4) and (5) for given values of the other parameters.[19]

1.3 Extensions

We discuss here several modeling extensions (see Web Appendix for further details). First, we consider an extension of the model where the household can make direct (costly) investments to reduce its concurrent propensity to use electricity. We show that the same three empirical objects as in equation (2) measure an upper bound for the deadweight loss and thus a lower bound for the bias (in absolute values). The inverse demand curve $p_1(x_1)$ will still capture the utility loss of changing x_1 to $\overline{x_1}$, taking into account such endogenous investments. Similarly, the inverse demand curve $p_2(x_2|\overline{x_1})$ will capture the utility loss of changing x_2 for the level of the propensity to consume resulting from changes in electricity use and investments in period 1. However, if the household had anticipated the policy in period 2, it might have made different investment choices in period 1, thus reducing its overall utility loss.[20] A similar argument would apply if we were to extend the model to allow for other endogenous behaviors such as saving and borrowing.

Second, we can allow for heterogeneity in households' preferences and in the degree of hysteresis. In this case, the social cost of a policy that aims at aggregate changes in behavior can be evaluated using aggregate inverse demand curves and the aggregate degree of hysteresis in equation (2), as long as the corrective policy is implemented efficiently (e.g., using tradable quotas or Pigouvian taxes).[21] In so doing, one would measure again an upper bound for the deadweight loss and a lower bound for the bias (in absolute values). We recognize that the temporary policy that we study may not have allocated the aggregate change in behavior efficiently across agents. However, we are not directly interested in the social cost of a long-run version of the exact same policy. We then abstract from allocative inefficiencies when incorporating our estimates in our framework to illustrate the importance of taking hysteresis into account for the social cost of corrective policies.

Finally, households could be "naive" about the effect of their current choices on their future propensity to consume. In this case, a corrective policy would also address an "internality" problem

[19] A linear approximation is likely less accurate for large changes in behavior, but our purpose is only to illustrate the possible magnitude of the bias.

[20] The investment cost function is assumed to be fixed. Instead, supply-side innovations in response to the policy (new energy-efficient appliances) could lower the cost of reducing one's propensity to use electricity. This would further reduce the deadweight loss (in absolute values). In our empirical setting, markets for domestic appliances are integrated across regions, so supply-side innovations are unlikely to explain our estimated long-term impact.

[21] An implication is that we can allow for the behavior of interest to change in discrete amounts at the individual level as long as the aggregate demand curves are smooth. This would be the case in a model with fixed adjustment costs as long as the thresholds triggering adjustments are smoothly distributed in the population.

by reducing consumption levels, which were suboptimally high from a private perspective. This would be another reason why neglecting the possibility of hysteresis overestimates the social cost of the policy. Agents could also have biased beliefs about the effect of their current choices on their future propensity to consume. As long as agents underestimate such an effect, this also would be a way in which neglecting the possibility of hysteresis overestimates the social cost of the policy. The opposite would be true if agents overestimate the effect. We provide some suggestive evidence that, if anything, agents underestimated this effect in our empirical application.

1.4 Connecting the theory to the data

A few comments must be made before moving to the empirical exercise.

First, the empirical analog of the key statistic $x_2(\overline{x_1}) - x_{20}$ is the *long-term* impact of a temporary corrective policy. This is important because short- and long-run elasticities are likely to differ. As a result, some persistence of an impact in the aftermath of a corrective policy may result from the slow convergence toward a prior steady state rather than from hysteresis. It is clear that the deadweight loss of a policy that aims at changing behaviors in the long run should be based on long-run elasticities, even in the absence of hysteresis. How long is "long enough" for the persistence to capture hysteresis will depend on the behavior under consideration and on whether the estimated persistence remains stable, rather than decaying over time.

Second, it is important to estimate the long-term impact of a policy that agents expected to be temporary. Otherwise, part of the persistence of a temporary policy may result from past (sunk) responses in anticipation of a possible continuation of the policy. This would not be a problem in the simple model presented above because the only way to anticipate a possible continuation of the policy is to change electricity use in period 1. However, it would be a problem in a model with, e.g., direct investments in the propensity to consume. In this case, past investments in anticipation of a possible continuation of the policy will affect future electricity use directly, not only through their effect on electricity use during the temporary policy. As a result, one would overestimate the effect on future behaviors of inducing changes in behaviors during the temporary policy.

Finally, several theories suggest that hysteresis is more likely for larger changes in behaviors.[22] Brief constraints may also be too short for agents to form new habits or to learn new ways to behave efficiently. Therefore, the ideal experiment to provide evidence of hysteresis and discuss implications for the social cost of long-run corrective policies would randomize a policy that aims at large changes in behavior, for a temporary period that is not too short and that agents know to be temporary, and whose impact can be estimated over a long period of time afterward.

[22] For instance, learning effects are more likely for policies that force agents to depart further from their usual levels of behavior. There may also be important fixed costs. Hysteresis in a rational habit formation model relies on pushing agents far enough from their prior steady state (Becker and Murphy, 1988).

2 Background and data

In the remainder of this paper, we assess the degree of hysteresis in a behavior of interest by exploiting a natural experiment that shares many features of the ideal experiment described above. We study the 10-year impact of a 9-month long policy in Brazil that was known to be temporary, and that led to the largest short-run reductions in household electricity use among temporary electricity saving programs around the world (Meier, 2005). In a natural experiment, identification requires additional assumptions, but it would be challenging for a controlled experiment to share all these features. One limitation of the policy that we study is that it did not rely on a single efficient and easily-replicable instrument to achieve its short-run reductions in electricity use. It included individually-assigned quotas and a set of (nonlinear) pecuniary and non-pecuniary incentives for households to consume below their quotas, which changed over time and may have affected households differently. We have little to say about the design of the specific policy in our empirical setting or its optimality. Our interest is that the policy allows us to provide clear evidence of hysteresis in a policy-relevant context. We then use our estimates to illustrate the possible implications of neglecting hysteresis for the social cost of long-run corrective policies.[23]

2.1 The temporary electricity saving program of the 2001 crisis

The temporary electricity saving program that we study was implemented in response to an exceptional shortage in electricity supply in specific areas of Brazil in 2001. Starting with a brief overview of the electricity distribution system, we provide here the necessary information about the 2001 crisis and its electricity saving program. More information is available in the Web Appendix.

2.1.1 Electricity distribution in Brazil

The major national electricity system in Brazil is divided into four subsystems: North (6.5% of total load in 2000), Northeast (14.5%), Southeast/Midwest (62%), and South (17%). Almost all households had access to the electricity grid in the South and the Southeast/Midwest in 2000. In contrast, the electricity grid was not fully developed in the two other subsystems at the time, but it developed rapidly in following years, thanks to strong support from the federal government (e.g., the program "Luz Para Todos"). The Brazilian electricity system relies almost exclusively on hydrological resources. In 2000, hydropower was responsible for 81% of the production capacity and

[23] In our back-of-the-envelope calculations, we must also assume that the social cost comes only from distorting quantities. This may or may not be the case, especially with non-pecuniary incentives, but such considerations are beyond the scope of our paper. Note that any policy that aims at changing aggregate behavior on such a large scale would likely stimulate a set of non-pecuniary incentives (e.g., peer pressure), at least endogenously.

94% of the electricity generated in the country (ONS, 2011). More than 60 local monopolies (distribution utilities) distribute electricity to end-consumers and housing units are typically metered and billed separately every month. Finally, electricity theft – i.e. illegal connections to the grid – is a serious concern in Brazil, amounting to 15% of the total load for some distribution utilities.

Electricity prices are regulated by a federal agency (Agência Nacional de Energia Elétrica, ANEEL). The main residential tariff is a flat unit price per kilowatt hour (kWh). An alternative tariff for low-income and small consumers offers percentage discounts on the main tariff depending on the quantity consumed. Price changes typically modify the main tariff and therefore imply a proportional change in every marginal price. The regulatory framework is a price-cap mechanism. Every four or five years, prices are *revised* to guarantee the economic viability of distribution utilities. However, demand risk falls entirely on the utilities between revision years. Yearly price *adjustments* only factor in changes in non-manageable costs, e.g. energy costs (ANEEL, 2005).[24]

2.1.2 History of the 2001 electricity crisis

The 2001 electricity crisis was due to supply factors. In particular, it was due to exceptionally low streamflow in the rivers that serve hydropower plants in specific areas of the country, combined with infrastructure constraints on generation and transmission capacity. Figure 1a displays the evolution of hydro-reservoirs' water levels in the Southeast/Midwest and in the South. We focus on the two largest subsystems in our analysis because of their similar development stage at the time (see next section). Water levels follow a seasonal pattern in the Southeast/Midwest, with heavy rain upstream of the reservoirs replenishing them during the austral summer. Levels were low in the two subsystems by 2000. In the Southeast/Midwest, they reached their lowest point in 40 years (for the season) in March 2001 because of exceptionally low summer rainfall.[25] In contrast, generous rain dissipated any risk of shortages in the South. The South could not transfer its excess supply to the other subsystems because of the limited transmission capacity across subsystems. Moreover, while growth in demand had never outpaced growth in projected demand in the years prior to 2001 (see Web Appendix), it outpaced growth in generation capacity. This was a nationwide issue and experts later concluded that the South would have faced a similar crisis if its hydropower plants had experienced a similar situation as in the Southeast/Midwest (Kelman, 2001).

By late April, it became clear that electricity use had to decrease to avoid generalized blackouts. The government announced that an incentive-based electricity saving program would start in June,

[24] The price-cap mechanism encourages distribution utilities to address electricity theft. Price revisions and adjustments occur at different times for different utilities. In June 2001, the main tariff was US$.23/kWh in Rio de Janeiro. Marginal prices in the alternative tariff were US$.08 (up to 30 kWh), US$.14 (up to 100 kWh), US$.20 (up to 140 kWh), and US$.23 (above 140 kWh). Monetary values are in US$ of 2012 throughout the paper (R$1.82=US$1). Minimum consumption levels are also charged, and local taxes increase what customers eventually pay.

[25] The Northeast and the North were facing a similar situation (not shown).

although details remained unclear (*O Globo*, April 23, 2001).[26] Distribution utilities supported instead the use of rolling blackouts because "financial penalties were unlikely to succeed, in part due to the lack of demand elasticity" and the expected length of the crisis (*Veja*, May 3, 2001; Maurer, Pereira and Rosenblatt, 2005). Rolling blackouts remained part of a plan B that never became necessary. The government program started on June 4, 2001, and, from the very start, it was expected to last until February 2002 (end of the next rainy season; *Veja*, July 19, 2001). The objective was to reduce electricity use by 20% in the Southeast/Midwest. The program also applied in the Northeast, and to a lesser extent in the North, but not in the South. Mation and Ferraz (2011) provide evidence that the crisis, the electricity saving program, and its differential implementation across subsystems were unanticipated, even by industrial customers.[27] As expected, the crisis officially ended in February 2002, but according to a specialized periodical, "people were giving signals that they learned how to avoid wasting electricity" (*Energia Elétrica*, March 15, 2002).

2.1.3 Incentives of the electricity saving program

The electricity saving program included individually-assigned quotas and a set of incentives for residential customers to consume below their quota. Rules were frequently repeated in the media and on electricity bills, but they were relatively complicated, nonlinear, and changed more than once.[28] Therefore, it is unclear exactly to which incentives individual households were responding. However, we are not interested in the specificities of these measures. Our interest is that households reduced electricity use dramatically in response to the policy, which was known to be temporary.

Every customer was assigned a *quota* at the start of the crisis. The typical quota was equal to 80% of a *baseline* corresponding to their average consumption from May to July 2000. Quotas for small consumers were set at 100% of baseline or 100 kWh, whichever was smaller.[29]

The incentive scheme included sticks and carrots. Customers exceeding their quota were charged a *fine* per kWh consumed above 200 kWh. Fines thus targeted larger consumers. The

[26]This was despite a first set of national policies in early April. Among these measures were the giveaway of efficient lightbulbs in low-income neighborhoods, a 15% reduction in electricity consumption in federal public buildings, the import of energy from Argentina, and the construction of new thermoelectric facilities (*Veja*, April 5, 2001).

[27]For instance, President Cardoso's approval rates dropped differentially in areas subject to the electricity saving program after its announcement. Figure 1b also shows that households in the different subsystems did not seem to have anticipated the electricity saving program by changing electricity use in different amounts prior to June 2001.

[28]Firms and the public sector also faced incentives in this period. Mation and Ferraz (2011) look at impacts on firms. We do not consider firms because the nature of their response to temporary corrective incentives may be very different and because changes in the industrial composition of the economy complicates the study of long-term effects. The fact that firms were subject to some incentives is an issue for our purpose only to the extent that it indirectly affected households' electricity use. However, we control for employment or income effects in our empirical analysis.

[29]Quotas were revised upward in December 2001 and January 2002 because the situation was improving. Consumption levels are also typically higher in the austral summer. Customers were informed of their quotas by mail prior to their first affected billing cycle. We reproduce such a letter in the Web Appendix. We also display the exact mapping between quotas and baseline consumption levels.

fine was equal to 50% of the usual tariff up to 500 kWh and then 200% of the tariff. Customers who exceeded their quotas more than once were also under the *threat of power cuts* of three to six days. Customers consuming less than their quota were eligible for a *bonus* per kWh reduced below their quota. The only guaranteed bonus was for customers consuming less than 100 kWh. The remaining funds from collecting fines would then be divided among other complying customers. The bonus payed per kWh saved below the quota was equal to at most 200% of the tariff.[30]

In practice, the implementation of these incentives was not smooth. First, households' response to the program was so large that fines did not raise enough money to pay any non-guaranteed bonus. The government then introduced another guaranteed bonus in September 2001 for customers with quotas below 225 kWh, with a unit bonus of 100% of the tariff. Second, distribution utilities did not have enough staff to implement power cuts. So, power cuts were limited to a very few customers and distribution utilities were asked to prioritize those who repeatedly consumed far above their quota. Power cuts were even prohibited in Rio de Janeiro (Lei Municipal 3266/2001), the city for which we have customer-level billing data.

Finally, a massive *conservation appeal* campaign (social incentives) was carried out in collaboration with distribution utilities and media outlets. Daily reports on TV compared achievements to policy targets. Energy conservation advice and stories of exemplary behavior were shared in the media to promote awareness and encourage participation. Media reports and messages on electricity bills included appeals to social preferences and patriotism. The government made sure to impose a more stringent conservation target for public buildings to set an example.

Households were also offered a lot of information in the media on how to reduce electricity use. This information reached the whole country, as the main media outlets are national in Brazil. It may thus explain some spillover effects to unaffected areas, but it is unlikely to explain any large differential trend between areas subject or not to the electricity saving program. Even if it explained part of the difference, our framework would apply as long as households internalized that information differently because they were subject to the policy. In fact, we would argue that any policy that aims at changing behavior on such a large scale would spur a market for information.

Other factors could have affected electricity use, but not differentially between the Southeast/Midwest and the South. Tariff changes followed the usual regulatory framework during and after the crisis.[31] Some policies were implemented nationally and are thus part of our counterfactual. Taxes on efficient lightbulbs were reduced, and taxes on electric showers, water heaters, and incandescent lightbulbs were temporarily increased (Decreto 3827, May 21, 2001). Efficiency

[30]Fines and bonuses were added into electricity bills, which could not be negative, limiting the payment of bonuses. We illustrate how these incentives modified the cost of electricity in the Web Appendix.

[31]This is with the exception of a 2.9% extraordinary increase for distribution utilities subject to the electricity saving program (*Camara de Gestão da Crise de Energia, Resolução 91*, December 21, 2001). Such a small price change is unlikely to drive any of our results. Moreover, we control for electricity tariffs in our empirical analysis.

standards for domestic appliances were adopted (Lei 10295, October 17, 2001), but only implemented in later years. Finally, the rainfall pattern during the 2000-2001 summer was a clear outlier, so there was no rational reason for customers in different subsystems to differentially update their beliefs about the risk of future shortages.[32] Even if they had changed their beliefs, it is unclear that households would have consumed less electricity in response, in light of the fact that the policy used grandfathering to assign individual quotas.

2.2 Data

Our analysis relies mostly on three sets of data.

A. Utility level administrative data (ANEEL).

Our main results are based on longitudinal data at the level of distribution utilities. The regulator (ANEEL) provided us with monthly administrative data from mandatory reports of distribution utilities on total electricity consumption, total revenues, and total number of customers by category (e.g., residential) from 1991 to 2011. Our main outcome, average residential electricity consumption per customer, is equal to total residential consumption divided by the number of residential customers. We also gathered copies of every tariff regulation published by the regulator from 1996 to 2011. As a result, we have a balanced panel of average residential electricity consumption and residential electricity tariff at the monthly level for all distribution utilities. In 2000, we have 26 utilities in the Southeast/Midwest and 17 in the South (numbers vary slightly with sample years due to some concession areas being split). We match our panel of distribution utilities to decennial census data (2000 and 2010) and to yearly data on population (1996-2011), GDP (1999-2011), formal employment, and average temperature (1996-2010), which are available at the municipality level and can be matched using information on the concession area of each distribution utility.[33]

B. Household level billing data (LIGHT).

We use longitudinal data at the customer level for one distribution utility subject to the electricity saving program to evaluate the robustness of our results and to go beyond average effects. We obtained individual billing data from January 2000 to December 2005 for the universe of low voltage customers of LIGHT, the distribution utility serving Rio de Janeiro and 31 surrounding municipalities in the Southeast. The data include about three million residential customers in 2000.

[32] Accordingly, when the government established an *insurance fund* to prevent subsequent crises, it chose to finance it through a nationwide undifferentiated increase in electricity tariffs ($.51 per 100 kWh; Camara de Gestão da Crise de Energia, Resolução 115). Reservoir levels were very low in 2000 even in the South. They were also more variable in the South after the crisis (see Figure 1a). Moreover, the country had already experienced smaller weather-induced electricity shortages in all subsystems in the past (Maurer, Pereira and Rosenblatt, 2005). These previous shortages led to isolated blackouts and not to the implementation of any incentive-based electricity saving program.

[33] Census, GDP and demographic data are from the Brazilian Institute of Geography and Statistics (IBGE). Formal employment data is from the Ministry of Labor (RAIS data). Temperature data is from Matsuura and Willmott (2012).

They detail every bill component and include metering and billing dates, meter location, and the quantity consumed in each month. Customers are uniquely identified over time unless they move.

C. Household level survey data (PPH).

We exploit the microdata from the two most recent rounds of the Survey of Appliances and Utilization Habits (PPH, Pesquisa de Posse de Equipamentos e Hábitos de Uso) to shed light on the mechanisms behind our results. The surveys were conducted by the National Electrical Energy Saving Program (PROCEL). A representative sample of residential customers from several utilities was surveyed before the crisis (first round, July 1998 to June 1999) and several years after the crisis (second round, July 2004 to June 2005). The in-house interviews included questions on household characteristics, appliance ownership, and consumption habits. Interviewers were asked to check some of the information directly, e.g., the number of lamps in the living room. We have a repeated cross-section of 8,803 households and 5,448 households from the same ten distribution utilities in 1999 and 2005, respectively. We have two utilities in the South (total of 3,121 households) and eight utilities in the Southeast/Midwest (total of 11,130 households).

D. Other data.

Finally, we confirm some of our results using time-series data on sales of appliances from manufacturers, as well as the Brazilian Household Expenditure Surveys (POF, Pesquisa de Orçamentos Familiares, with rounds in 1996-1997, 2002-2003, and 2008-2009). Those data are not available at the municipality level and cannot be matched to the concession areas of distribution utilities.

3 Empirical strategy and descriptive statistics

Our main empirical strategy exploits the panel of distribution utilities through a difference-in-difference, comparing utilities in the Southeast/Midwest and in the South over time. In this section, we first provide descriptive statistics supporting our approach. We then present our main empirical specification and a series of tests to evaluate the robustness of our results.

3.1 Descriptive statistics

Table 1 provides some descriptive statistics supporting our key identification assumption of a common trend in average residential electricity consumption for distribution utilities in the Southeast/Midwest and in the South. It also shows that such an assumption is unlikely to hold, especially in the long run, when considering distribution utilities in the other two subsystems subject to the electricity saving program. Columns (1)-(5) compare initial values across distribution utilities in the four subsystems (and for LIGHT separately), namely the mean and range of relevant variables in 2000. Columns (6)-(8) present the differential trend in these variables between 2000 and 2010, comparing utilities in each of the three other subsystems to utilities in the South.[34] Some of the variables discussed below are presented in a similar table in the Appendix.

Before comparing subsystems, note that average residential electricity use was lower in Brazil (below 200 kWh/month in 2000) than in more developed countries (903 kWh/month in the US in 2012; www.eia.gov/tools/faqs). Households were of course poorer and less likely to own major domestic appliances, but electricity was also relatively expensive. The main residential electricity tariff in 2000 (US$.18/kWh) was higher than the US average price in 2012 (US$0.12/kWh).

There are two main reasons not to consider distribution utilities from the Northeast and North subsystems in our analysis. First, nearly all households in the South had electricity prior to the crisis. This was not the case in the Northeast and in the North. Moreover, the share of households with electricity increased substantially in these subsystems in the following decade, by 8%-11% compared to the South. A common-trend assumption is unlikely to hold when customer bases evolve very differently. Second, households in the Northeast and in the North were poorer and less likely to own major domestic appliances prior to the crisis – e.g., we observe no common support for the median household income and the share of households owning refrigerators between utilities in the Northeast and in the South. Strong poverty alleviation, as experienced in Brazil in the 2000s, can have very different effects on residential electricity use when initial ownership rates are so different (Wolfram, Shelef and Gertler, 2012). Accordingly, the share of households owning a refrigerator, a major domestic appliance in terms of electricity use, increased by 15%-26% in the North and in the Northeast compared to the South in the following decade. This is unlikely to be a consequence of the crisis and thus clearly violates a common-trend assumption.[35]

In contrast, distribution utilities in the South constitute a more credible counterfactual for distribution utilities in the Southeast/Midwest. First, in both subsystems, nearly all households had electricity prior to the crisis. Second, customer bases evolved at a similar rate in the following decade. Moreover, this does not hide any differential trend in population size, access to electricity, urbanization, household size, dwelling size, or dwelling characteristics. Third, initial ownership rates of major domestic appliances (refrigerator, washing machine, TV, air conditioner) and subsequent growth in ownership rates are comparable. Fourth, initial electricity tariffs were also comparable

[34] Specifically, we use yearly data from the most recent census years ($t \in 2000, 2010$) and the following specification:

$$log\left(y_{d,t}\right) = \alpha_d + \gamma\, \mathbb{1}\{t = 2010\} + \delta\, \mathbb{1}\{t = 2010\, \&\, d \in \text{TreatRegion}_d\} + \nu_{d,t}$$

where a_d is a fixed effect for distribution utility d, and *TreatRegion* indicates a distribution utility from a subsystem subject to the electricity saving program. $\nu_{d,t}$ is an error term clustered by utility. Columns (6)-(8) report estimates of $\widehat{\delta}$ for samples including utilities from one of the other three subsystems and the South.

[35] Additionally, there is a data limitation preventing us fro studying outcomes in the North subsystem. Many customers there are served by isolated electricity systems. Our utility-level panel data do not differentiate residential consumption from "isolated" and "connected" customers, and the former were not subject to the electricity saving program. The policy also started later (August) and ended earlier (December) in the North.

and did not evolve differentially. If anything, tariffs decreased relatively in the Southeast/Midwest.

Distribution utilities in the South and in the Southeast/Midwest do, however, differ in some respects. For instance, average electricity use and median income levels were on average higher in the Southeast/Midwest in 2000. Median income levels also grew relatively less in the Southeast/Midwest in the following decade, even though labor market outcomes such as employment, formal employment, or farm employment did not evolve differentially (we attribute the differential trend in average electricity use to the impact of the electricity saving program below). Importantly, the range of initial values overlapped for all these variables and many others, as did the range of changes in these values in subsequent years (see Web Appendix). Therefore, we can control for many relevant variables without relying on purely parametric assumptions (this would not be true using data from the Northeast). One factor that differs systematically between the Southeast/Midwest and the South is climate. Average temperatures are higher in the Southeast/Midwest. This is unlikely to drive any of our results. We show in the Web Appendix that there is no relationship between changes in average residential electricity use in the Southeast/Midwest and in the South, and either levels of average temperatures or changes in average temperatures. In sum, the data appear to support our common-trend assumption, as shown in Figure 1b for earlier periods, at least conditionally on controlling for some relevant factors.

3.2 Empirical strategy

We estimate the short- and long-term impacts of the temporary electricity saving program through a generalized difference-in-difference. We regress the logarithm (or the level) of average residential electricity consumption per customer for utility d from region r in calendar month m of year t using the following specification:

$$
\begin{aligned}
log\left(Av\_kWh_{d,r,m,t}\right) =& \alpha_d + \beta_{r,m} + \gamma_p \mathbb{1}\left\{t \in TimePeriod_p\right\} \\
&+ \delta_p\, \mathbb{1}\left\{d \in \text{SE/MW}\ \&\ t \in TimePeriod_p\right\} + log\left(X_{d,r,m,t}\right) + \nu_{d,r,m,t} \qquad (6)
\end{aligned}
$$

where the coefficients α_d and $\beta_{r,m}$ are fixed effects for each distribution utility and for each calendar month per region (seasonality), respectively. We divide our monthly observations into various time periods, indexed by p. We consider yearly time periods before and after the crisis. We divide 2001 and 2002 into three time periods: pre-crisis (early 2001), crisis (June 2001–February 2002), and post-crisis (rest of 2002). The coefficients γ_p are time-period fixed effects. The coefficients δ_p then capture difference-in-difference estimators in each time period for the impact of the temporary electricity saving program. Finally, we cluster error terms, $\nu_{d,r,m,t}$, at the level of the distribution utility, and distribution utilities are weighted equally in our regressions.[36]

Estimates of δ_p during the crisis and for subsequent time periods can be interpreted as the average treatment effect of the policy on the treated under a common-trend assumption. Of course, it is not possible to test whether this assumption holds in practice. However, we show in Figure 1b that average electricity use had been following roughly the same trend since 1991 in the South and the Southeast/Midwest. Estimates of δ_p for time periods preceding the crisis will directly test for the presence of a common-trend prior to the crisis. Moreover, as explained in Section 2, the timing of the crisis and the differential treatment between subsystem was entirely due to supply factors that are likely exogenous to potential changes in households' electricity use. A common-trend assumption is thus reasonable in our context, especially in the short run. As we are particularly interested in long-run effects, we then reinforce the common-trend assumption by controlling for variables that may be correlated with other factors affecting electricity use. In particular, we use the following variables available at the distribution utility level in each year ($X_{d,r,m,t}$): main residential electricity tariff (1996-2011), population size (1996-2011), GDP (1999-2011), formal employment levels, and average temperature (1996-2010). Our preferred specification includes all these controls; therefore, the sample is restricted to a balanced panel of utilities between 1999 and 2010.

3.3 Robustness checks

We undertake four types of empirical tests to evaluate the robustness of our results. First, we obtain similar results when estimating variants of equation (6) with different sample years and the controls available in those years. Second, we show that our estimated average impact is not driven by outliers. To do so, we separately estimate the impact of the electricity saving program for each distribution utility using synthetic control methods (Abadie, Diamond and Hainmueller, 2010). The synthetic control estimator is the difference between an outcome in a *treated* utility during and after the crisis, and the same outcome in a "synthetic" weighted average of *control* utility. The outcome of interest, $Y_{d,t}$, is the demeaned seasonally adjusted logarithm of average monthly residential consumption. The vector of weights W is chosen to minimize: $\| Y_{d0} - Y_{c0}W \| = \sqrt{(Y_{d0} - Y_{c0}W)' V (Y_{d0} - Y_{c0}W)}$, where Y_{d0} and Y_{c0} are vectors containing the values of the outcome in pre-crisis periods in the treated utility and in control utilities, respectively. An optimal choice of V minimizes the mean squared error of the synthetic control estimator.

Third, we investigate whether our estimated long-term impact is robust to controlling for variables that are not available at the yearly level but that are available in the 2000 and 2010 censuses

[36]Our independence assumption seems reasonable. The two subsystems cover large and heterogeneous areas, providing electricity to more than 100 million individuals. Moreover, electricity sector policy is centralized at the federal level, including policies related to efficiency. In the Web Appendix, we show similar results weighting distribution utilities by their customer base. Those results are noisier as they are driven by the few very large distribution utilities.

($X_{d,t}$), such as median household income. We show graphically the relationship between changes in these variables and changes in average electricity consumption between 2000 and 2010. We also estimate the following regression using only data from census years:

$$log(Av\_kWh_{d,t}) = \alpha_d + \gamma\, \mathbb{1}(t = 2010) + \delta\, \mathbb{1}(t = 2010\,\&\, d \in \text{SE/MW}) + log\left(X_{d,t}\right) + \nu_{d,t} \quad (7)$$

where $\nu_{d,t}$ is an error term for utility d in census year t clustered by distribution utility.

Fourth, we use the longitudinal data for LIGHT customers to address a series of confounding explanations that cannot be tackled with utility-level data. We compare changes in average electricity use in LIGHT aggregate monthly data and in a random sample of individual customers observed in every month from 2000 to 2005. This balanced panel does not suffer from composition effects by construction. We then document the persistence of changes in consumption levels during the crisis *at the individual level*.[37] Finally, we investigate other dimensions of the distribution of changes in consumption levels, which are interesting in themselves and also as a way to address a concern that "electricity theft" responses may explain some of our results.

4 Results

We first present results using utility-level data. We then turn to the customer-level data. Most of our regression results are presented graphically, but corresponding tables with coefficient estimates and standard errors are in the Web Appendix.

4.1 Utility-level data

Difference-in-difference coefficients $\widehat{\delta_p}$ from estimating our preferred specification of equation (6) are displayed in Figure 3 with their 95% confidence intervals. The sample is restricted to the balanced panel of utilities between 1999 and 2010 so that we can include all the controls available at the yearly level. Panel (a) and (b) consider specifications in logs and levels, respectively. The first few months of 2001 are used as a reference period. Point estimates are close to 0 prior to the crisis, supporting our common-trend assumption. Average residential electricity consumption then dropped differentially in the Southeast/Midwest when the electricity saving program came into force. We estimate an impact of -.26 log points (23%) during the crisis, or 41.5 kWh/month.

[37] We cannot estimate the causal effect of some policy variation among LIGHT customers. First, all customers were subject to some aspects of the policy and the rules were complicated and uncertain. Second, we don't know how non-pecuniary incentives were perceived by different customers. Third, there is no discontinuity or bunching in electricity use at the quota or at other levels where customers faced discontinuities or kinks in their budget. This is consistent with the existing literature (Borenstein, 2009). In an earlier working paper, we showed that even large quasi-exogenous variation in quotas led to only very small variation in consumption (Gerard, 2013).

This is the first quasi-experimental estimate of the short-run impact for residential customers, but it is already well known that electricity consumption was successfully reduced at the time. Our main contribution is to show that about half of the short-run impact persisted in the long run. Consumption levels partially rebounded after the crisis but point estimates are stable from 2005 onward at about -.115 log points (11%), or 19 kWh/month. It is thus unlikely that our estimates are due to a difference between short- and long-run elasticities rather than to hysteresis.

Our results are similar if we consider all the possible combinations of sample years determined by the availability of our yearly controls: 1991-2011 (no controls), 1996-2011 (tariffs and population), 1996-2010 (tariff, population, formal employment, and average temperature), and 1999-2011 (tariff, population, and GDP). Point estimates for the log specification are displayed in Figure 4a. We omit confidence intervals for the sake of clarity. The estimated short- and long-run impacts are almost identical in all specifications. Moreover, Figure 4a shows that average electricity consumption followed similar trends in the Southeast/Midwest and in the South since at least 1991. Point estimates are slightly positive between 1997 and 2000, a pattern that is apparent in the raw data in Figure 1b, but they get closer to 0 as we add more controls. Results are also similar if: (i) we consider winter and summer months separately, (ii) if we weight distribution utilities by their customer base at baseline, and (iii) if we restrict the sample to distribution utilities with overlapping support in average electricity use (outcome) and in household median income (not available at the yearly level) at baseline to reinforce our common-trend assumption (see Web Appendix).

Figure 4b displays synthetic control estimates of the impact of the electricity saving program for each distribution utility. Monthly estimates are averaged into the same time periods as in Figure 3. Darker lines correspond to utilities in the Southeast/Midwest. Lighter lines correspond to placebo estimates in which we compare a given distribution utility in the South to a weighted average of the others. The synthetic controls are able to closely match the trends pre-crisis, including the period between 1997 and 2000. The estimated short-run impact is large for all the distribution utilities subject to the policy: between -.19 and -.40 log points. Importantly, the long-run impact is also negative for all those utilities. Our estimated average impact is thus not driven by outliers (this can be seen in the raw data in the Web Appendix). The median and the average of the utility-specific impacts are in fact comparable – e.g. -.13 and -.14 log points in 2011, respectively. In contrast, the median and the average of our placebo estimates are very close to 0 in all years.[38]

Table 2 displays estimates of the long-term impact of the electricity saving program that are obtained by using equation (7) and controlling for variables available in the 2000 and 2010 censuses. Columns (4)-(6) restrict the sample to distribution utilities with overlapping support at the baseline in average electricity use and household median income. The estimated impact is similar when we

[38]We replicated this exercise for the main electricity tariff. In that case, the range of estimated impacts overlaps completely and is centered around 0 for utilities in the Southeast/Midwest and in the South (see Web Appendix).

don't include any control (columns 1 and 4), when we control for the main electricity tariff and median household income (columns 2 and 5), and when we add controls for population size, the share of households living in urban areas, average household size, average dwelling size, the share of dwellings with a bathroom, the employment rate, and the average temperature (columns 3 and 6). The robustness of our results does not come from an absence of variation in these variables. We show graphically in the Web Appendix that long-term changes in consumption levels are systematically lower for utilities in the Southeast/Midwest than in the South for given baseline levels or long-term changes in all the variables in Table 1 and in its continuation table in the Web Appendix.

4.2 Customer-level data

Figure 5 presents robustness checks using the longitudinal microdata for LIGHT customers. Panel (a) shows that the time-series in average electricity use for LIGHT is similar when we use (i) the aggregate data at the utility level, (ii) microdata from a random sample of customers in each month, and (iii) microdata from a random sample of customers observed in every month from 2000 to 2005 (balanced panel). This provides additional evidence that composition effects, absent from the balanced panel by construction, are unlikely to drive our results, at least until 2005 (the LIGHT-specific impact remains large after 2005 in Figure 4b).[39] It also implies that our estimates are not due to electricity theft at the extensive margin. Customers who are obtaining all their electricity through illegal connections to the grid at some point are excluded from the balanced panel.[40]

Panel (b) shows that average changes in electricity use came from sizable reductions at every level of consumption. It displays kernel densities for average monthly electricity use in our balanced panel before, during, and after the crisis. The density during the crisis is stochastically dominated. Densities one year and four years after the crisis are similar and fall between the crisis and pre-crisis densities. During the crisis, 92% of customers consumed less electricity than before the crisis; the median customer reduced electricity use by 31%. Four years after the crisis, 69% were still consuming less electricity; the median customer was consuming 16.5% less electricity.[41]

Panel (c) shows that average changes in electricity use came from large reductions from most

[39]There is no evidence of economically meaningful migration across regions between 2000 and 2010. de Oliveira and de Oliveira (2011) document that the Southeast experienced a net out-migration during the period. However, magnitudes are very small: no larger than 0.2% of the Southeast population and 0.5% of the South population.

[40]There is no good data on electricity theft. Distribution utilities are supposed to report yearly information on distribution losses to the regulator, but many did not provide this information prior to 2000. In the Web Appendix, we use yearly reports for 24 utilities in the Southeast/Midwest and in the South from 1998 to 2008. The data are very noisy and, if anything, point estimates suggest that non-technical losses (a measure of theft) decreased compared to 2000. In the Web Appendix, we also use microdata from the Brazilian Household Budget Survey (POF 1996/97, 2002/03 and 2008/09) and find no differential trend in the share of households who do not pay for electricity.

[41]We show similar patterns for summer and winter months separately in the Web Appendix. We show that customers with higher baseline consumption levels made larger proportional reductions in electricity use. By considering shifts in the distribution of electricity use in Figure 5b, we avoid mean reversion issues (Borenstein, 2009; Ito, 2014).

customers *at a given baseline consumption level*. It displays the distribution of changes in average monthly electricity use during and after the crisis, compared to before the crisis, for a balanced panel of randomly selected customers with the same baseline consumption (around 300 kWh). Kernel densities are based on electricity use during the first five months of the crisis, and in the same months in other years, when these customers faced the same quota and incentives. As mentioned before, we find no evidence of bunching at the quota. During the crisis, 98% of customers reduced electricity use and the median customer reduced usage by 34%. Four years after the crisis, 78% were still using less electricity than before the crisis; the median customer was using 22% less electricity. We show similar patterns for other baseline consumption levels in the Web Appendix.

Together, panels (b) and (c) also provide evidence against electricity theft responses at the intensive margin. Establishing illegal connections to the grid to obtain part of one's own electricity use is more likely to take place among poorer and smaller consumers. However, reductions in electricity use were not concentrated among small consumers. We show in the Web Appendix that they were also large among customers from relatively wealthy neighborhoods. Finally, if electricity theft takes place among relatively large consumers, it is likely to be concentrated within a small group. This is inconsistent with the evidence in panel (c) that most customers at a relatively large baseline consumption level severely reduced their electricity use during the crisis.

So far, we have implicitly interpreted our results as evidence of hysteresis at the individual level. Panel (d) shows that there is indeed a strong correlation between electricity use during the crisis and four years after the crisis for customers with the same baseline consumption levels (using the same sample as in panel c). In the Web Appendix, we show similar patterns for other baseline consumption levels. We provide evidence on the mechanisms of hysteresis in the next section.

5 Mechanisms of hysteresis

There are three main mechanisms that could explain a long-run reduction in household electricity use. Households may have permanently changed the quantity of appliances that they own, the type of appliances that they own, or their utilization of these appliances. We shed some light on each of these mechanisms below. To do so, we rely primarily on household-level microdata from the two most recent rounds of the Appliances and Utilization Habits surveys (PPH, 1998-1999 and 2004-2005). We use data from ten distribution utilities – eight in the Southeast/Midwest and two in the South – that were surveyed in both rounds.[42] Conveniently, our estimated long-term impact on household electricity use is stable starting around the time of the second PPH survey round (see Figure 3). We also use other data sources that are described in more detail in the Web Appendix.

[42] We were not given the identity of those distribution utilities due to confidentiality concerns. Therefore, we cannot match the PPH survey data to the ANEEL administrative data.

5.1 Asking households directly

A natural starting point to investigate the mechanisms of persistence is to ask households directly. The second round of PPH surveys included a special section for customers of distribution utilities subject to the electricity saving program during the 2001 crisis. For each major domestic appliance, it asked households whether: (1) they were using the appliance as much as before the crisis; (2) they were using it less than before the crisis; (3) they had disconnected or disposed of the appliance during or after the crisis; (4) they had substituted a more energy-efficient model during or after the crisis. Households could choose only one answer and we display the share that chose each answer in Table 3, panel B. For each appliance, we also show in panel A the average quantity per household and the estimated average monthly electricity use for customers of these utilities prior to the crisis.

A large share of households who owned a given appliance prior to the crisis reported using it less after the crisis. This is true for electric showers (39%), lights (41%), and freezers (21%),[43] which together account for a significant portion of electricity use in Brazil, but also for other appliances. The one exception is for refrigerators, another main source of electricity use. This is in fact reassuring as we don't expect much flexibility in refrigerators' utilization. In contrast, less than 3% of households reported replacing appliances with more energy-efficient models, except for lights (9%). Finally, a large share of households reported disconnecting or disposing of their appliance only in the case of freezers (17%) and air conditioners (5%).

The main stated mechanism of persistence is a change in utilization habits. However, households may have changed their utilization habits *and* their appliance stock. Households in the South may also have made some similar changes. We thus further investigate the three mechanisms.

5.2 Appliances' quantity

The PPH surveys recorded data on the quantity of a list of appliances for households in the Southeast/Midwest and in the South in both survey rounds. We investigate any differential trend in the quantity of appliances using a difference-in-difference strategy as in the previous sections:

$$Y_{h,d,t} = \alpha_d + \gamma\, \mathbb{1}\,(t = 2005) + \delta\, \mathbb{1}(t = 2005 \,\&\, d \in \text{SE/MW}) + log\left(X_{h,d,t}\right) + \nu_{h,d,t} \qquad (8)$$

where $Y_{h,d,t}$ is an outcome for household h from distribution utility d in survey round t. We control for utility fixed effects α_d and a survey round fixed effect γ. The coefficient δ is a difference-in-difference estimator under a common-trend assumption. We cannot provide evidence of a common

[43] An electric shower is an electric heating device placed in a shower head. It is a popular technology in developing countries where gas is primarily available through gas tanks. It has a low fixed cost but consumes a lot of electricity. It is common in many countries for households to have smaller refrigerators than in the US, with a small freezer unit, but to have a separate larger (horizontal or vertical) freezer unit.

trend prior to the crisis with two repeated cross-sections. We thus control for household characteristics, $X_{h,d,t}$, which may be correlated with different trends in appliance ownership.[44] We also construct an appliance quantity index to avoid multiple-inference problems, normalizing the quantity of each appliance using the average and standard deviation of appliance ownership in the South in 1999 (Kling, Liebman and Katz, 2007). We display difference-in-difference estimates in Table 4, panel A, for the five main domestic appliances in terms of electricity use. Results for other appliances are in the Web Appendix. Standard errors are obtained using the wild cluster bootstrap-t, given our small number of clusters (Cameron, Gelbach and Miller, 2008). The resulting confidence intervals are large, typically including 0, so our results based on PPH data remain suggestive.

Point estimates are negative for our index and for all appliances, except for lights. They are close to 0 for refrigerators and washing machines, which is consistent with findings based on census data (see Section 3.1). They are large in magnitude for freezers and air conditioners, consistent with the information reported in Table 3. Finally, the coefficient is large in magnitude and significant for TVs. Using census data, we found no difference in the share of households with a TV (see Section 3.1). PPH measures instead the number of TVs per household.[45]

5.3 Appliances' characteristics

The PPH surveys recorded some appliance characteristics correlated with electricity use. We use the specification in equation (8) to investigate any differential trend in these characteristics and in two indices, one for the age of appliances and one for the type (size/power). The sign of all variables is normalized, such that a positive sign implies a higher propensity to use electricity. Results are displayed in Table 4, panel B, for the same five main domestic appliances and in the Web Appendix for other appliances.

Point estimates are positive for our "age" index (older) and for the age of each of our main domestic appliances. Replacing appliances with newer models, which likely consume less electricity, may have been difficult for Brazilian households, who are relatively poorer and face a much higher cost of credit than in more advanced countries.[46] In fact, the supply side of the market

[44]The vector of household characteristics include income, squared income, number of household members, dwelling size, and dummies identifying wealthier neighborhoods and neighborhoods close to slums ("favelas").

[45]The difference between the two results may also be due to a difference in sampling (the census is representative of the Brazilian population; PPH is representative of customers of the ten unidentified utilities who are regularly connected to the electricity grid) or to a difference in timing (post-crisis data is from 2005 in the PPH survey and from 2010 in the census) – e.g. a temporary effect if, as we discuss in the next section, sales of appliances decreased during the crisis. In the Web Appendix, we obtain similar estimates using microdata from the Household Budget Survey (POF) from 1996/97, 2002/03 and 2008/09. By 2008/09, we find roughly zero point estimates for the quantity of refrigerators and TVs, and a negative (and larger in magnitude) point estimate for the quantity of freezers owned by households.

[46]In 2001, Brazil was the country with the highest real interest rate in the World Development Indicators of the World Bank. It was 44.65 percent, compared to an average of 8.34 percent for OECD countries.

for domestic appliances ex-ante expected, and ex-post reported, losses from the electricity crisis (*Folha de São Paulo*, June 5, 2001 and March 6, 2002). In the Web Appendix, we show that there is no discontinuous increase in estimates of national monthly sales of major domestic appliances during the crisis. In fact, there is a discontinuous decrease in sales for many appliances.[47]

At the margin, we would still expect households to prefer models that consume less electricity when buying an appliance during the crisis. Point estimates are negative for the size of our main domestic appliances, although standard errors are again large. In the Web Appendix, we show some related evidence for electric showers: the average power of electric showers sold by a leading Brazilian manufacturer decreased differentially in the Southeast/Midwest during the crisis (by about 10%), but it increased again after the crisis.[48]

Finally, we showed in Table 3 that households reported adopting more energy-efficient lightbulbs in the Southeast/Midwest. It is well known that compact fluorescent lightbulbs (CFLs) spread rapidly in Brazil during and after the crisis. This national pattern is present in the PPH data; we estimate an average increase of 52 percentage points in the share of CFLs in the South and Southeast/Midwest between survey rounds. However, the difference-in-difference estimate suggests that adoption rates were higher in the South than in the Southeast/Midwest. As a result, the coefficient for our "type" index is positive despite the negative coefficients for the other appliances.

5.4 Utilization habits

The PPH surveys recorded utilization habits correlated with electricity use for many appliances. As above, we use the specification in equation (8) to investigate any differential trend in each of these habits and in a "utilization habit" index. The signs of all variables are again normalized such that a positive sign implies a higher propensity to use electricity. Results are displayed in Table 4, Panel C, for the same five appliances and in the Web Appendix for other appliances.

Point estimates are large and negative for our index and for utilization habits related to our main domestic appliances, again except for lights. For instance, households were much less likely to have their separate freezer unit permanently switched on in the Southeast/Midwest after the crisis, which is consistent with information in Table 3. Households were also less likely to set the thermostat of their electric shower on the warmest setting; this result is statistically significant even using our large standard errors. A back-of-the-envelope calculation suggests that this behavior

[47] We obtained those data from Whirlpool, a leading manufacturer, which produces those estimates for its market strategy. It did not share with us the estimation methodology used. At the time of the crisis, the Southeast/Midwest accounted for more than 55% of the national market of refrigerators, for example, and the South for about 20%.

[48] We obtained data on the monthly sale of all the models of electric showers from Fame, a leading manufacturer, disaggregated by state. The data include the power (wattage) of each model, which is the only relevant measure of electric showers' propensity to use electricity. We could not find similar data for other domestic appliances. The PPH data is too noisy to look at the characteristics of appliances of different ages.

alone could have generated enough savings to explain the long-term impact (22 kWh/month).[49]

In sum, the main stated mechanism of hysteresis is a change in habits (living with fewer appliances is also a new habit). We found suggestive evidence that households indeed changed their habits in the Southeast/Midwest. However, we cannot reject the possibility that some changes in appliances' characteristics also played a role. Note that our conceptual framework applies independently of the specific mechanisms as long as the hysteresis was due to the temporary policy.[50]

6 Implications for the social cost of corrective policies

We have established that households reduced electricity use in the long run in response to a temporary electricity saving program. In this section, we use our evidence to illustrate the extent to which neglecting such hysteresis could overestimate the social cost of long-run corrective policies.

First, a temporary policy delivers a persistent correction of the externality in the presence of hysteresis if the prior level of the behavior was too high from a social perspective. The associated gains could balance out part of the social cost of the policy. We estimated that the temporary policy reduced electricity use by 11% in the Southeast/Midwest until at least 2011, which corresponds to 67.8 billion kWh. PROCEL, the electrical energy saving program of the federal government, claims to have obtained reductions of 39.6 billion kWh from 2002 to 2011 at a discounted cost of US$380 million (5% yearly discount rate). At that average level of cost efficiency, PROCEL would have had to spend US$650 million to achieve a similar reduction in electricity use. It is unclear how PROCEL measures its impact and our calculation is thus likely to be a lower bound.

Second, we evaluate equations (4) and (5) to illustrate the implication of the estimated degree of hysteresis when calculating the social cost of a long-run corrective policy. Consider a policy aimed at reducing electricity use by 23% (average effect during the crisis) over a period of 10 years. Period 1 corresponds to the nine months of the temporary policy and period 2 to the following 111 months. We estimated a necessary residual reduction in electricity use in the second period of 23%-11%=12%. In the Web Appendix, we estimate a medium-run price elasticity of $-.291$ for average residential electricity use, using yearly variation in electricity tariffs across distribution utilities after the crisis. Applying this estimate, the change in marginal utility caused by the policy

[49] The thermostat of an electric shower can be switched off or set at either "Low Power" (*Modo Verão*) or "High Power" (*Modo Inverno*). An electric shower consumes on average 30% less electricity in Low Power than in High Power. Our kWh figure is obtained by multiplying the estimated impact (-.863), the gain from setting the shower to Low Power (30%) and the average electricity consumption of electric showers in High Power (87.1 kWh/month).

[50] In the second round of PPH surveys (but not in the first round), households were also asked about the information that they had learned about energy efficiency. In the Web Appendix, we show that households in the Southeast/Midwest were more likely to have obtained information on energy efficiency labels and energy savings behaviors after the crisis, but less likely to know what that information implies for how much electricity they could actually save. Note that any energy efficiency policy in Brazil, including information provision, would be implemented nationally by PROCEL.

in period 2 is equivalent to a 79% increase in electricity tariffs, assuming away hysteresis, but only a 41.2% increase given our estimated degree of hysteresis. The bias in level is then:

$$|DWL_{NoH}| - |DWL| = \frac{1}{2}\sum_{t=10}^{120}\beta^t \frac{p_t X_t/.89}{.2911}\left[(-.23)^2 - (-.12)^2 \frac{1}{.89/1}\right] = US\$4.14 \text{ billions}$$

where we use the same discount rate as above, and where $p_t X_t$ is the total monthly bill for residential electricity use in the Southeast/Midwest after the crisis. It is scaled by $\frac{1}{.89}$ to obtain counterfactual amounts ($p_{20}x_{20}$). A bias of US\$5.82 billion corresponds to 6.59% of the total (discounted) monthly bill for residential electricity use in the Southeast/Midwest after the crisis. Even if the long-run elasticity was three times larger than the medium-run elasticity, the bias would still amount to US\$1.94 billion. Another way to provide some perspective is to express the bias as a percentage of the deadweight loss. Using equation (5), we have:

$$\frac{|DWL_{NoH}| - |DWL|}{|DWL|} = \frac{\sum_{t=10}^{t=120}\beta^t \frac{p_t X_t}{.89}\left[(-.23)^2 - (-.12)^2 \frac{1}{.89/1}\right]}{\sum_{t=1}^{t=9}\beta^t \frac{p_t X_t}{.89}(-.23)^2 + \sum_{t=10}^{t=120}\beta^t \frac{p_t X_t}{.89^2}(-.12)^2} = 1.82$$

Assuming away hysteresis, one could then overestimate the social cost of the policy by 182%. This is just an illustrative calibration, but it shows how neglecting the possibility of hysteresis could severely overestimate the social cost of long-run corrective policies. Moreover, the overestimate would be even higher if we assumed a longer time horizon.

So far, we have assumed that households perceived the true costs and benefits of changing their behavior. These costs must have been relatively high, given that our estimates imply that an average household could permanently save 11% of its electricity bill by reducing electricity use by 23% for a nine-month period. This contrasts with households' reported experience during the crisis. When asked about changes to their quality of life during the crisis in PPH surveys, only 24% of households in the Southeast/Midwest reported a decrease in quality of life. Most households reported no change in quality of life (48%) or that they had "learned to live comfortably while saving money" (28%, see Table A.1 in the Appendix). Moreover, most households who managed to reduce electricity below their quota reported that it was either "not difficult at all" (43%) or that it was "not very difficult" (48%). The contrast between large long-run private gains and limited short-run private costs suggests that, if anything, before the crisis households were underestimating the returns from changing their behavior. This echoes several theories and findings in the literature. For instance, Bryan, Chowdhury and Mobarak (2014) find that it is difficult to explain the low level of seasonal urban migration in their context given the high returns of a one-time incentivized migration and the persistent re-migration rate. Agents with hyperbolic time discounting would underinvest in their habit formation, even if they were aware of their habit formation process (Gruber and Köszegi, 2001). Of course, agents may not be aware of such a process (Acland and Levy, 2015). Finally, households may have correctly understood the private costs of changing their behavior. However, these costs may drop dramatically when everybody is subject to a corrective policy because of, e.g., social learning (Dupas, 2014). The contrast between large long-run private gains and limited short-run private costs of changing behavior is also consistent with the so-called energy paradox (Jaffe and Stavins, 1994).[51]

7 Conclusion

This paper argues that one might overestimate the social cost of long-run corrective policies by neglecting the possibility of hysteresis in the behavior of interest. We provided evidence of the importance of hysteresis in a policy-relevant behavior. We then used our estimates in a simple conceptual framework to illustrate the possibly large magnitude of such a bias.

A limitation of our empirical analysis is that we cannot relate the estimated impact to one single easily-replicable policy instrument, and policymakers may not be interested in permanently implementing a policy similar to the temporary policy that we studied. Yet, there is a lot of interest in reducing the demand for resources, such as energy and water. A main argument against implementing sizable corrective policies in that context is that the associated social cost could be too large. Our paper implies that such an argument may be overstated. The big question that this study then leaves open is *how* to trigger the emergence of energy-efficient habits in a credit-constrained, low-income, and low-consumption setting. Shaping consumption patterns at an earlier stage of development may help attenuate the pressure of energy demand growth in developing countries.

The application in this paper considers the context of energy demand. However, our argument applies to any context where hysteresis may be relevant. For instance, Fujiwara, Meng and Vogl (2014) find that a weather-induced 1 percentage point rise in turnout in an election increased voter turnout in the following election by 0.9 percentage points. Assuming that this impact persists in the long run, one could thus overestimate the social cost of long-run corrective policies by neglecting the degree of hysteresis in turnout. Using their estimates and similar back-of-the-envelope assumptions as in our Section 6, the social cost of a policy that aims at a 1 percentage point permanent increase in voter turnout could be overestimated by 1450%.[52]

[51]This also could be explained in a rational inattention model. Attention costs for energy-saving opportunities, rather than the direct cost of changing behavior, may have been high before the crisis (Sallee, 2014). Such costs would be factored into the demand curve for electricity use and our framework would thus apply.

[52]We assume a constant cost of voting, a constant baseline turnout level (at their mean of 58%), a yearly discount rate of 5%, an infinite time horizon, and a constant elasticity with respect to voting costs.

References

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2010. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association*, 105(490): 493–505.

Acemoglu, Daron, Philippe Aghion, Leonardo Bursztyn, and David Hemous. 2012. "The Environment and Directed Technical Change." *American Economic Review*, 102(1): 131–166.

Acland, Dan, and Matthew R Levy. 2015. "Naiveté, projection bias, and habit formation in gym attendance." *Management Science*, 61(1): 146–160.

Allcott, Hunt. forthcoming. "Site selection bias in program evaluation." *Quarterly Journal of Economics*.

Allcott, Hunt, Allan Collard-Wexler, and Stephen D O'Connell. 2014. "How do electricity shortages affect productivity? Evidence from India." *NBER Working Paper 19977*.

Allcott, Hunt, and Dmitry Taubinsky. forthcoming. "Evaluating Behaviorally–Motivated Policy: Experimental Evidence from the Lightbulb Market." *American Economic Review*.

Allcott, Hunt, and Todd Rogers. 2014. "The Short–Run and Long–Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation." *American Economic Review*, 104(10): 3003–3037.

ANEEL. 2005. "Tarifas de fornecimento de energia elétrica." *Agência Nacional de Energia Elétrica, Cadernos Temáticos ANEEL*, 4.

Bardelin, Cesar. 2004. "Os efeitos do racionamento de energia elétrica ocorrido no Brasil em 2001 e 2002 com ênfase no consumo de energia elétrica." *Dissertação de Mestrado em Sistemas de Potência, Universidade de São Paulo: Escola Politécnica*.

Becker, Gary, and Kevin Murphy. 1988. "A Theory of Rational Addiction." *Journal of Political Economy*, 96(4): 675–700.

Borenstein, Severin. 2009. "To What Electricity Price Do Consumers Respond? Residential Demand Elasticity Under Increasing-Block Pricing." *Mimeo, Energy Institute, University of California at Berkeley*.

Bryan, Gharad, Shyamal Chowdhury, and Ahmed Mushfiq Mobarak. 2014. "Underinvestment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh." *Econometrica*, 82(5): 1671–1748.

Cameron, C., J. Gelbach, and D. Miller. 2008. "Bootstrap-Based Improvements for Inference With Clustererd Errors." *Review of Economics and Statistics*, 90(3): 414–427.

Charness, Gary, and Uri Gneezy. 2009. "Incentives to Exercise." *Econometrica*, 77(3): 909–931.

Davis, Lucas W, Alan Fuchs, and Paul Gertler. 2014. "Cash for coolers: evaluating a large–scale appliance replacement program in Mexico." *American Economic Journal: Economic Policy*, 6(4): 207–238.

de Oliveira, Luiz Antônio Pinto, and Antônio Tadeu Ribeiro de Oliveira. 2011. *Reflexões sobre os deslocamentos populacionais no Brasil.* IBGE.

Dixit, Avinash. 1989. "Entry and Exit Decisions under Uncertainty." *Journal of Political Economy*, 97: 620–638.

Dupas, Pascaline. 2014. "Short-Run Subsidies and Long-Run Adoption of New Health Products: Evidence From a Field Experiment." *Econometrica*, 82(1): 197–228.

Ferraro, Paul J, and Michael K Price. 2013. "Using nonpecuniary strategies to influence behavior: evidence from a large-scale field experiment." *Review of Economics and Statistics*, 95(1): 64–73.

Fisher-Vanden, Karen, Erin T Mansur, and Qiong Juliana Wang. 2015. "Electricity Shortages and Firm Productivity: Evidence from China's Industrial Firms." *Journal of Development Economics*.

Fujiwara, Thomas, Kyle Meng, and Tom Vogl. 2014. "Estimating Habit Formation in Voting." *Mimeo, Princeton University.*

Gerard, François. 2013. "What Changes Energy Consumption, and for How Long? New Evidence from the 2001 Brazilian Electricity Crisis." *RFF Discussion Paper 13–06, E3 WP-050*.

Giné, Xavier, Dean Karlan, and Jonathan Zinman. 2010. "Put your money where your butt is: a commitment contract for smoking cessation." *American Economic Journal: Applied Economics*, 213–235.

Gruber, Jonathan, and Botond Köszegi. 2001. "Is Addiction "Rational"? Theory and Evidence." *The Quaterly Journal of Economics*, 116(4): 1261–1303.

Hintermann, Beat, and Andreas Lange. 2013. "Learning abatement costs: On the dynamics of the optimal regulation of experience goods." *Journal of Environmental Economics and Management*, 66(3): 625–638.

IPCC. 2014. "Climate Change 2014: IPCC Fifth Assessment Report."

Ito, Koichiro. 2014. "Do Consumers Respond to Marginal or Average Price? Evidence from Nonlinear Electricity Pricing." *American Economic Review*, 104(2): 537–563.

Ito, K., T. Ida, and M. Tanaka. 2015. "The Persistence of Moral Suasion and Economic Incentives: Field Experimental Evidence From Energy Demand." *NBER Working Paper*, 20910.

Jaffe, Adam, and Robert Stavins. 1994. "The energy paradox and the diffusion of conservation technology'." *Resource and Energy Economics*, 16(2): 91–122.

Jessoe, Katrina, and David Rapson. 2014. "Knowledge is (Less) Power: Experimental Evidence from Residential Energy Use." *American Economic Review*, 104(4): 1417–1438.

Kelman, Jerson. 2001. "Desequilíbrio entre Oferta e Demanda de Energia Elétrica." *Relatório da Comissão de Análise do Sistema Hidrotérmico de Energia Elétrica*.

Kling, J., J. Liebman, and L. Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica*, 75(1): 83–119.

Levinson, Arik. 2014. "California energy efficiency: Lessons for the rest of the world, or not?" *Journal of Economic Behavior and Organization*, 107: 269–289.

Mation, Lucas, and Claudio Ferraz. 2011. "How Do Firms React to Infrastructure Constraints? Evidence from Brazil's Energy Shortage." *Working paper, Department of Economics, Pontifícia Universidade Católica do Rio de Janeiro*.

Matsuura, Kenji, and Cort Willmott. 2012. "Terrestrial precipitation: 1900-2010 gridded monthly time series (1900-2010)(v 3.01 added 6/14/12). University of Delaware." *http://climate.geog.udel.edu/ climate/*.

Maurer, Luiz, Mario Pereira, and José Rosenblatt. 2005. "Implementing Power Rationing in a Sensible Way: Lessons Learned and International Best Practices." *Energy Sector Management Assistance Program Report, World Bank*.

McKinsey. 2009. "Pathways to a Low–Carbon Economy: Version 2 of the Global Greenhouse Gas Abatement Cost Curve." *Report from McKinsey & Company*.

Meier, Alan. 2005. "Saving Electricity in a Hurry: Dealing with Temporary Shortfalls in Electricity Supplies." *OECD/IEA*.

Miller, Conrad. 2014. "The Persistent Effect of Temporary Affirmative Action." *Mimeo, Massachussets Institute of Technology*.

ONS. 2011. "Plano Decenal de Energia." *Tech. report, Operador Nacional do Sistema Eletrico*.

Pigou, Arthur. 1920. *The Economics of Welfare.* London: McMillan & Co.

Pimenta, Amanda, Hilton Notini, and Luiz Felipe Maciel. 2009. "Brazilian Electricity Demand Estimation: What Has Changed after the Rationing in 2001? An Application of Time Varying Parameter Error Correction Model." *Mimeo, Getulio Vargas Foundation*.

Reiss, Peter, and Matthew White. 2008. "What Changes Energy Consumption? Prices and Public Pressures." *RAND Journal of Economics*, 39(3): 636–663.

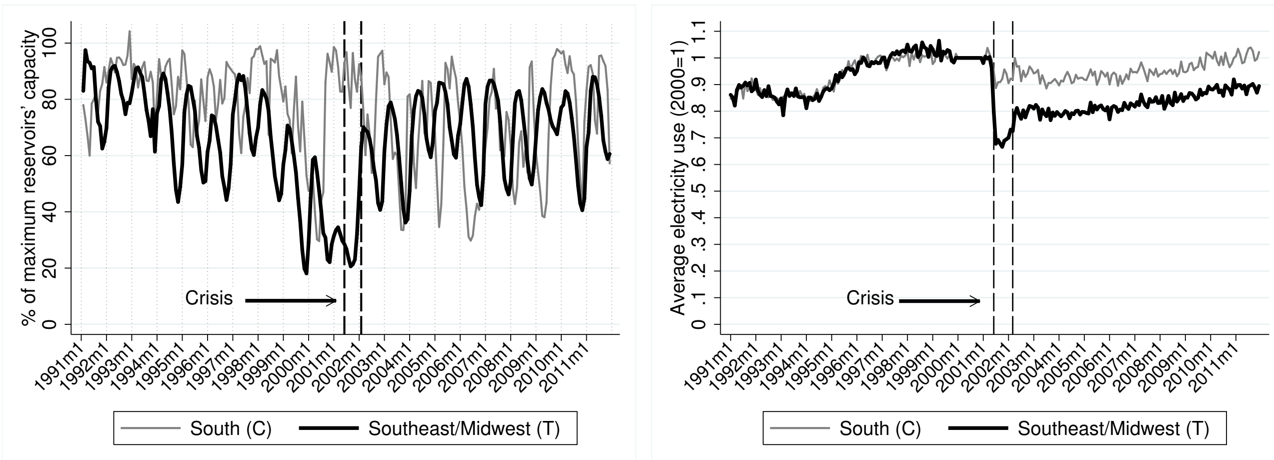
Sallee, James. 2014. "Rational Inattention and Energy Efficiency." *Journal of Law and Economics*, 57(3): 781–820.

Wolfram, Catherine, Orie Shelef, and Paul Gertler. 2012. "How Will Energy Demand Develop in the Developing World?" *Journal of Economic Perspectives*, 26: 119–138.

Figure 1: Causes and consequences of the temporary electricity saving program

(a) Stocked energy of the hydroelectric reservoirs

(b) Average residential electricity use

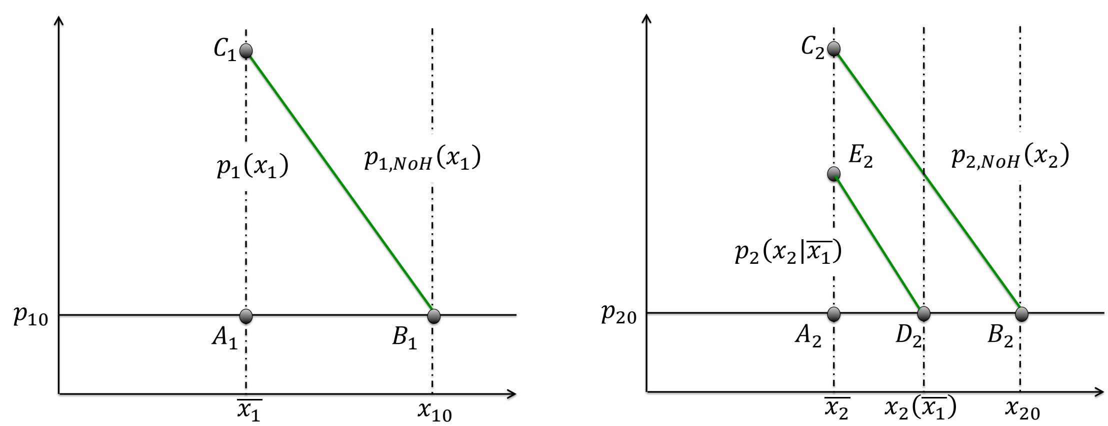


Panel (a) displays the evolution of hydroelectric reservoirs' capacity in the Southeast/Midwest and in the South (dotted lines indicate January; official data from ONS, the National System Operator). In the summer of 2000–2001, exceptionally low rainfall upstream of the reservoirs led to dangerously low levels in the reservoirs in the Southeast/Midwest. In order to prevent supply shortages, an electricity saving program was implemented from June 2001 to February 2002 (dashed lines). No such program was implemented in the South, where generous rain dissipated any risk of shortages. Panel (b) displays the overall impact of the electricity saving program on monthly average residential electricity consumption per customer for distribution utilities in the two subsystems (utility-level administrative data). We present unweighted averages in each month, normalized with respect to the same month in 2000 (seasonality). Trends were similar prior to June 2001. Average residential consumption then dropped, especially for distribution utilities in the Southeast/Midwest subject to the electricity saving program. It rebounded after February 2002, once the program was suspended, but only partially. Comparing patterns in the Southeast/Midwest and in the South suggests that the temporary electricity saving program had a stable and persistent impact on average residential consumption, until at least the end of our sample (2011).

Figure 2: The social cost (deadweight loss) of corrective policies with hysteresis

(a) Period 1

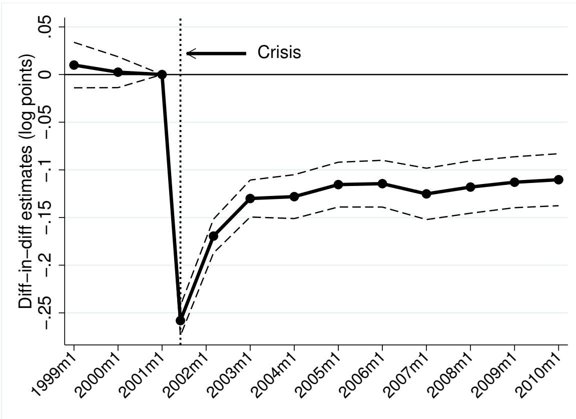
(b) Period 2



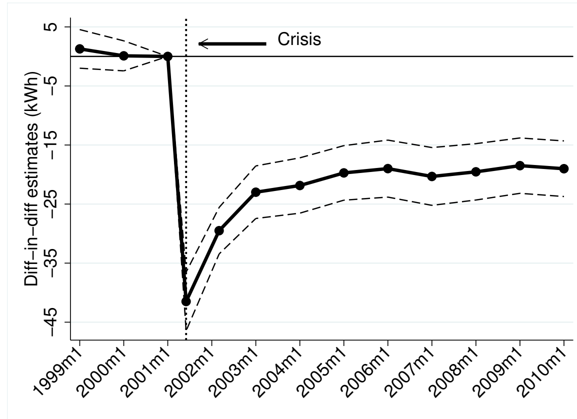
The figure illustrates how one could overestimate the social cost or deadweight loss of long-run corrective policies by neglecting the possibility of hysteresis. Without government intervention, x_{10} and x_{20} are the equilibrium quantities in period 1 and 2, given baseline prices p_{10} and p_{20}, respectively. Consider a corrective policy reducing quantities to $\overline{x_1}$ and $\overline{x_2}$. Reducing quantity in period 1 to $\overline{x_1}$ increases the household's marginal utility for x_1, which can be traced along the inverse demand curve $p_1(x_1)$. In the presence of hysteresis, this would reduce the quantity in period 2 to $x_2(\overline{x_1})$, absent any government intervention in period 2. The demand curve $p_1(x_1)$ would factor in any effect on utility and behavior in period 2 from changes in x_1. The loss in utility from reducing quantity in period 1 to $\overline{x_1}$ is then the triangle $A_1B_1C_1$, assuming locally linear demand curves. Holding constant $\overline{x_1}$, further reducing quantity to $\overline{x_2}$ in period 2 similarly increases the household's marginal utility for x_2 which can be traced along the inverse demand curve $p_2(x_2|\overline{x_1})$. The associated loss in utility is the triangle $A_2D_2E_2$. Neglecting the possibility of hysteresis (but assuming no bias from having the "wrong" shapes for the demand curves, such that the slopes of the demand curves when assuming away hysteresis – indexed by *NoH*– are similar), one would overestimate the deadweight loss in the second period. Tracing the change in marginal utility along the whole interval from x_{20} to $\overline{x_2}$, one would obtain a loss in utility corresponding to the larger triangle $A_2B_2C_2$. The bias is the area $D_2B_2C_2E_2$. An estimate of the degree of hysteresis, $x_2(\overline{x_1}) - x_{20}$, is key to evaluating this bias. It could be identified from the long-run impact of a corrective policy implemented only in period 1 (temporary policy).

Figure 3: Main difference–in–difference results for average residential electricity use

(a) in logs



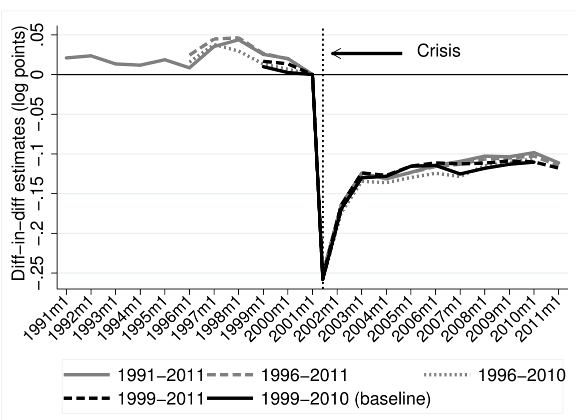
(b) in levels



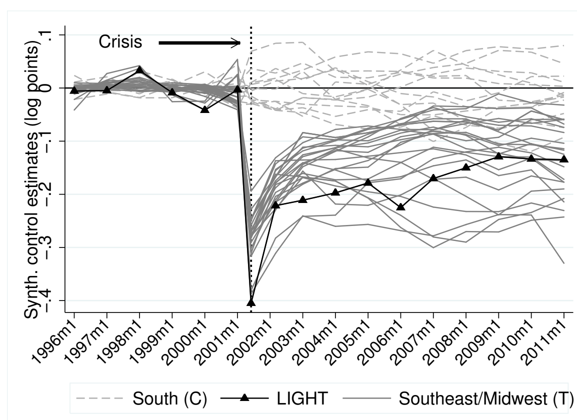
Utility-level administrative data for distribution utilities in the Southeast/Midwest and in the South from 1999 to 2010. The figures display coefficients from regressing the logarithm (panel a) or the level (panel b) of monthly average electricity consumption per customer for each utility on time–period dummies (yearly dummies, three dummies for 2001-2002 to isolate the crisis period), interacted with an indicator for utilities subject to the electricity saving program during the crisis (difference–in–difference estimators in each time period; 95% confidence interval in dashes, standard errors clustered by utility). The reference period corresponds to the first months of 2001. Regressions include uninteracted time–period dummies, utility and calendar month–per–region fixed effects, and controls for the logarithm or level of the main residential electricity tariff and of available yearly data matched to the concession area of each utility (population size, GDP, formal employment levels, average temperature). Point estimates are close to 0 prior to the crisis, supporting our common-trend assumption. Average residential electricity use then dropped differentially in the Southeast/Midwest when the electricity saving program came into force. We estimate an impact of -.26 log point (-23%) during the crisis, or 41.5 kWh/month. Consumption levels partially rebounded after the crisis but point estimates are stable from 2005 onward at about -.115 log points (-11%), or 19 kWh/month.

Figure 4: Robustness checks using utility–level data

(a) Difference–in–difference results for different sample years and the available controls (in logs)



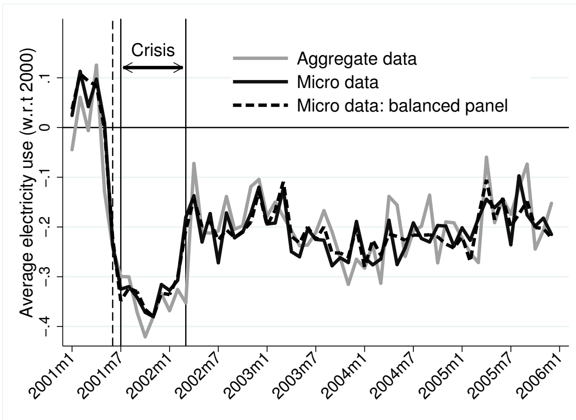
(b) Synthetic control estimates of utility–specific impact (in logs)



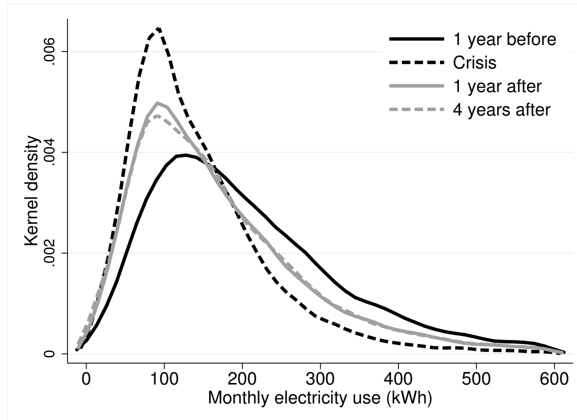
Panel (a) shows that our difference–in–difference estimates in Figure 3a are almost identical if we consider all the possible combinations of sample years determined by the availability of our yearly controls: 1991-2011 (no controls), 1996-2011 (tariffs and population), 1996-2010 (tariff, population, formal employment, average temperature), 1999-2011 (tariff, population, GDP). We omit confidence intervals for the sake of clarity (available in tables in the Web Appendix). This also shows that trends were similar in the Southeast/Midwest and in the South since at least 1991. Point estimates are slightly positive between 1997 and 2000, a pattern that is apparent in the raw data in Figure 1b, but they get closer to 0 as we add more controls. Panel (b) displays synthetic control estimates of the impact of the electricity saving program for each distribution utility. Monthly estimates are averaged into the same time periods as in Figure 3a. Darker lines correspond to distribution utilities in the Southeast/Midwest. Lighter lines correspond to placebo estimates in which we compare a given distribution utility in the South to a weighted average of the others. The synthetic controls are able to closely match the pre-crisis trends, including between 1997 and 2000. The estimated short-run impact is large for all the distribution utilities subject to the electricity saving program, between -.19 and -.40 log points. Importantly, the long-run impact is also negative for all those utilities. Our estimated average impact is thus not driven by outliers. The median and the average of the utility-specific impact is in fact comparable — e.g. $-.13$ and $-.14$ log points in 2011, respectively. In contrast, the median and the average of our placebo estimates are very close to 0 in all years.

Figure 5: Lessons and robustness checks using longitudinal microdata for LIGHT customers

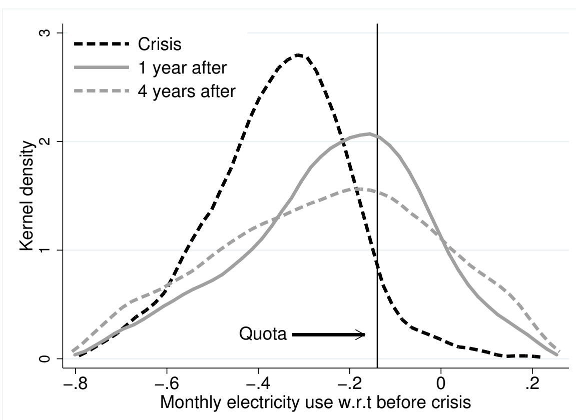
(a) Comparing time–series in average electricity using the utility–level data and the microdata



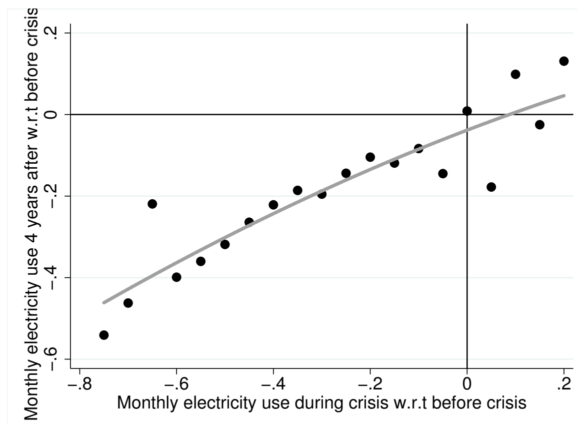
(b) Distribution of monthly electricity use over time for a balanced panel of customers



(c) Distribution of changes in monthly electricity use for a balanced panel of customers with the same baseline consumption level / quota



(d) Correlation between changes in electricity use during and after the crisis for a balanced panel of customers with same baseline consumption level / quota



Individual monthly billing data for the universe of residential customers of LIGHT (Southeast) from January 2000 to December 2005. Panel (a) displays average electricity use for LIGHT customers in each month compared to the same month in 2000. It shows that the time–series is almost identical when (i) we use the aggregate data at the utility level, (ii) microdata from a 2% random sample of customers in each month, and (iii) microdata from a random sample of customers observed in every month from 2000 to 2005 (balanced panel; 44,817 customers). This provides additional evidence that composition effects, absent from the balanced panel by construction, are unlikely to drive our results, at least until 2005. Panel (b) shows that average changes in electricity use came from sizable reductions at every level of consumption. It uses the same balanced panel to investigate changes in the distribution of electricity use over time. It displays kernel densities for monthly electricity use before, during, and after the crisis. Kernel densities are based on data from June to December, such that we can compare consumption levels up to four years after the crisis. The density during the crisis is stochastically dominated by the other densities. Densities one year and four years after the crisis are very similar and fall exactly between the crisis and pre-crisis densities. Panel (c) shows that average changes in electricity use came from large reductions from most customers at a given baseline consumption level. It displays the distribution of changes in electricity use during and after the crisis compared to the same months before the crisis, for a subset of the sample in panel (b) in which customers had about the same baseline consumption levels for quota assignment, and thus faced the same pecuniary incentives during the crisis (10% above and below 300 kWh/month; 4,344 customers). Kernel densities are based on electricity use during the first five months of the crisis (and in the same months in other years), before any change in quotas. The quota (vertical line) is not at -20% because it was based on the baseline months in 2000 (May to July) and not on these five months. We find no evidence of bunching at the quota. During the crisis, 98% reduced electricity use and the median customer reduced usage by 34%. Four years after the crisis, 78% were still using less electricity than before the crisis; the median customer was using 22% less electricity (mean 19%). Panel (d) displays the correlation between individual changes in electricity use during and after the crisis compared to the same months before the crisis, for the same sample as in panel (c). Customers are averaged by bins of 5% changes in electricity use during the crisis. The strong correlation suggests that the long-term impact is due to the persistence of individual changes in electricity use. We show similar patterns as in panels (c) and (d) for other baseline consumption levels in the Web Appendix. Kernel densities use Epanechnikov kernels and optimal bandwidths.

Table 1: Descriptive statistics for distribution utilities in the four subsystems

	Descriptive statistics in 2000 Mean [min-max]					Differential trends 2010 vs. 2000 Coefficient, in logs (s.e.)		
	South	LIGHT	South–East/Midwest	North–East	North	SE/MW vs. S	NE vs. S	N vs. S
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Average residential electricity	166	226	190	107	180	-.113***	-.001	.006
consumption (kWh/month)	[133–191]		[122–261]	[72.7–129]	[128–267]	(.021)	(.038)	(.048)
Main residential electricity	.175	.211	.189	.175	.173	-.095	-.003	-.084
tariff (US$/kWh)	[.154–.197]		[.162–.217]	[.162–.198]	[.148–.19]	(.061)	(.065)	(.098)
Number of customers	360	2864	861	834	238	.014	.151***	.165***
(1000's)	[1.63–2200]		[17.7–4160]	[63–2405]	[12.2–857]	(.041)	(.044)	(.045)
Population size	1457	9025	3235	4323	1605	.031	.027	.164***
(1000's)	[12.1–9208]		[64.8–16661]	[193–13014]	[124–6122]	(.03)	(.028)	(.047)
Share of households	.982	.999	.984	.894	.83	-.002	.081***	.109***
with electricity	[.949–1]		[.896–1]	[.759–.989]	[.645–.989]	(.007)	(.028)	(.037)
Share of households	.795	.991	.861	.72	.734	-.02	-.01	-.003
in urban areas	[.587–.916]		[.654–.991]	[.613–.851]	[.414–.979]	(.017)	(.017)	(.025)
Median household income	737	945	762	342	496	-.104**	.081*	-.057
(US$/month)	[508–945]		[450–1181]	[282–413]	[354–803]	(.045)	(.043)	(.082)
Share of households	.937	.972	.921	.646	.708	.009	.258***	.148***
with refrigerator	[.828–.994]		[.806–.985]	[.538–.758]	[.522–.916]	(.018)	(.037)	(.056)
Share of households	.436	.554	.366	.089	.199	-.008	.369***	.06
with washing machine	[.145–.661]		[.1–.618]	[.039–.137]	[.075–.311]	(.089)	(.092)	(.118)
Average temperature	18	21.9	21.5	25	26	-.032*	-.025	-.016
(degrees Celsius)	[16.7–19.5]		[19–24.5]	[23.1–26.4]	[25.2–26.5]	(.019)	(.02)	(.023)
Observations	17	1	26	11	8	86	56	50

Utility-level administrative data for distribution utilities in the four subsystems in 2000 and 2010 and census data matched to the concession area of these utilities in the same years; temperature data is from Matsuura and Willmott (2012). Columns (1)–(5) display descriptive statistics in 2000 (prior to the crisis) for the variables listed in the left–hand side column for distribution utilities in the South (column 1), in the Southeast/Midwest (LIGHT in column 2, all distribution utilities in column 3), in the Northeast (column 4), and in the North (column 5). Columns (6)–(8) display estimates of a long–term difference–in–difference estimator comparing the logarithm of these variables in 2010 vs. 2000 for distribution utilities in the Southeast/Midwest (column 6), in the Northeast (column 7), and in the North (column 8) compared to distribution utilities in the South. Significance levels: *10%, **5%, ***1% (s.e. clustered by utility). Regressions include fixed effects for each distribution utility and each year. All monetary values are expressed in US$2012. A similar table with additional variables is provided in the Appendix. We argue in the text that the information in this table shows that distribution utilities in the South constitute a credible control group for distribution utilities in the Southeast/Midwest, but not for those in the other two subsystems.

Table 2: Long–term difference in difference results (2010 vs. 2000)

Dependent variable: Log(yearly average household electricity consumption)						
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment region	-.113***	-.118***	-.103***	-.121***	-.119***	-.115***
× Year2010	(.021)	(.026)	(.028)	(.026)	(.029)	(.041)
Log main		-.201***	-.151		-.214**	-.166
residential tariff		(.071)	(.096)		(.089)	(.111)
Log median		.143	.318***		.151	.435**
household income		(.095)	(.118)		(.126)	(.171)
Clusters	86	86	86	70	70	70
Restricted sample	No	No	No	Yes	Yes	Yes
Other controls	No	No	Yes	No	No	Yes

Utility–level administrative data for distribution utilities in the Southeast/Midwest and in the South in 2000 and 2010 and census data matched to the concession area of these utilities in the same years; temperature data is from Matsuura and Willmott (2012). Significance levels: *10%, **5%, ***1% (s.e. clustered by utility). The table displays estimates of the long–term impact of the electricity saving program, controlling for census data. The first row displays coefficients from regressing the logarithm of monthly average electricity consumption per customer for each utility on a year dummy for 2010 interacted with an indicator for utilities subject to the electricity saving program during the crisis (long–term difference–in–difference estimators). All regressions include the uninteracted year dummy and utility fixed effects. Columns (4)–(6) restricts the sample to distribution utilities with overlapping support at baseline in average electricity use and in household median income. The estimated impact is similar when we don't include any control (columns 1 and 4), when we control for the main electricity tariff and median household income (columns 2 and 5), and when we additionally control for the logarithm of population size, the share of households living in urban areas, average household size, average dwelling size, the share of dwellings with a bathroom, the employment rate, and the average temperature (columns 3 and 6). The robustness of our results does not come from an absence of variation in these variables. We show graphically in the Appendix that long–term changes in consumption levels are systematically lower for utilities in the Southeast/Midwest than in the South for given baseline levels or long–term changes in all the variables in Table 1.

Table 3: Self-reported appliance usage after the crisis (Southeast/Midwest)

	Main domestic appliances					Other domestic appliances		
	Electric shower	Refrigerator	Freezer	Light	TV	Air Conditioner	Washing machine	Microwave
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. 1999 Survey								
Average (Quantity)	.97	.99	.20	8.45	1.39	.10	.53	.21
Average (kWh/month)	58.14	41.71	7.88	42.54	15.63	2.78	3.37	2.97
Panel B. 2005 Survey								
Share of households who owned appliance	.92	.97	.18	1	.96	.06	.68	.33
Conditional on prior ownership, share of households who:								
Use appliance as much as before the crisis	.61	.90	.59	.5		.30	.44	.53
Use appliance less than before the crisis	.39	.07	.20	.41		.56	.54	.39
Disconnected or disposed of the appliance	.01	0	.16	0		.05	.01	.03
Substituted a more energy-efficient model	0	.03	.01	.08		.03	0	0

Household-level data for eight distribution utilities in the Southeast/Midwest from the two most recent rounds of the Survey of Appliances and Utilization Habits (PPH, 1998/1999 and 2004/2005). Panel A displays the average number per household for the appliances listed on top of each column and the inputted average monthly electricity use in 1999, before the crisis. The latter is calculated by multiplying quantity by average utilization in 1999 (share of appliances owned frequently in use) and by the kWh consumption of the average model of each appliance from PROCEL estimates – shown in Table Q.1 in the Web Appendix. N=6482. Panel B reports the share of households who owned one of these appliances at some point in time and, in 2005, gave one of four possible answers for each appliance. The possible answers were: (1) households were currently using the appliance as much as before the crisis; (2) they were using it less than before the crisis; (2) they had disconnected or disposed of the appliance during or after the crisis; or (4) they had substituted a more energy-efficient model during or after the crisis. N=4579.

Table 4: Difference-in-difference results for appliances' quantity, characteristics, and utilization

Panel A. Quantity

	Index (KKL)	Shower	Refrigerator	Freezer	Light	TV
	(1)	(2)	(3)	(4)	(5)	(6)
$SE/MW \times Year2005$	**-.183**	**-.075**	**-.032**	**-.184**	**.669**	**-.329****
	(.248)	(.276)	(.042)	(.241)	(1.063)	(.153)
Average SE/MW 1999	-.020	.969	.994	.202	8.447	1.392
N	14,251	14,251	14,251	14,251	14,251	14,251

Panel B. Characteristics

	Index		Refrigerator		Freezer		Light		TV	
	Age (KKL)	Type (KKL)	Age	Size (Liters)	Age	Size (Liters)	CFLs (share)	Wattage (incand.)	Age	Size (Inches)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$SE/MW \times Year2005$	**.091**	**.034**	**.107**	**-38.373**	**.789**	**-2.706**	**-.263**	**-1.977**	**.230**	**-1.455**
	(.146)	(.111)	(.916)	(56.625)	(.728)	(28.849)	(.548)	(14.839)	(1.292)	(4.976)
Average SE/MW 1999	-.058	.198	7.501	304.793	5.193	239.519	.150	63.215	5.291	18.754
N	14,206	14,206	12,787	8,815	2,390	2,179	13,038	13,050	12,110	13,603

Panel C. Utilization

	Index (KKL)	Shower Thermostat	Appliance Always Switched On		Appliance Frequently Used	
		High Power	Refrigerator	Freezer	Light	TV
	(1)	(2)	(3)	(4)	(5)	(6)
$SE/MW \times Year2005$	**-.951**	**-.863****	**-.042**	**-.221**	**.595**	**-.473**
	(1.305)	(.425)	(.077)	(.270)	(1.401)	(.690)
Average SE/MW 1999	.212	.391	.973	.183	3.565	1.116
N	14,251	14,251	14,251	14,251	14,251	14,251

Household-level data for eight distribution utilities in the Southeast/Midwest and two in the South from the two most recent rounds of the Survey of Appliances and Utilization Habits (PPH, 1998/1999 and 2004/2005). The table displays difference-in-differences estimates of the impact of the energy saving program on the quantity, characteristics, and utilization of the five main electrical appliances, from equation (8) in Section 5. Each estimate corresponds to a regression of a different dependent variable and appliance (listed on top of each column). Panel A considers the quantity of appliances owned by households, Panel B the indicated characteristics of appliances owned, and Panel C the quantity of appliances frequently used or the quantity of electric showers regulated on high power (winter mode). The *KKL indices* consider the average of the dependent variables in the other columns, each normalized by their average and standard deviation in the South in 1999 (Kling, Liebman and Katz, 2007). For these indices, we input missing values with the mean of the cell the household belongs to (South/Southeast-Midwest and 1999/2005). All regressions include utility fixed effects, year fixed effects, income, squared income, number of household members, floorplan area, and dummies indicating a wealthy neighborhood and proximity to a slum. We input missing values in two control variables (income and dwelling size), using a linear regression, for each year, of the variable on the household's level (class) of energy use and the remaining controls. Significance levels: *10%, **5%, ***1% (s.e. estimated with the wild-cluster bootstrap-t by utility level). We show similar results for other available domestic appliances in the Web Appendix.

A Appendix Tables

Table A.1: Households' reported life quality during the crisis (Southeast/Midwest)

	Percentage of respondents
	(1)
How do you evaluate your change in life quality caused by the electricity saving program?	(N=4376)
I did not experience any change in life quality (*não houve variação*)	.48
I experienced some discomfort (*causou desconforto*)	.20
I experienced some severe discomfort (*causou muito desconforto*)	.08
I learned to live with the same comfort while saving money (*aprendi a viver com o mesmo conforto economizando dinheiro*)	.24
If your consumption reduction was sufficient to meet your quota, how difficult was it?	(N=3375)
It was very difficult (*muito*)	.09
It was not so difficult (*pouco*)	.48
It was not difficult at all (*nenhuma*)	.43

Household-level data for eight distribution utilities in the Southeast/Midwest from the two most recent rounds of the Survey of Appliances and Utilization Habits (PPH, 1998/1999 and 2004/2005). The table displays the percentage of households who answered each of the two questions indicated. These were not open questions and households had to choose one of the answers or "Others". In Portuguese, the two questions were "*Como o(a) sr.(a) avalia a variação de qualidade de vida causada pelo racionamento?*" and "*As medidas adotadas para atingir as metas durante o período de racionamento foram suficientes ou mais que suficientes. Dificuldade?*". The answers in Portuguese are displayed next to our translation.

Table A.2: Additional descriptive statistics for distribution utilities in different subsystems

	Descriptive statistics in 2000 Mean [min-max]					Differential trends 2010 vs. 2000 Coefficient, in logs (s.e.)		
	South	LIGHT	South–East/Midwest	North–East	North	SE/MW vs. S	NE vs. S	N vs. S
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Average residential electricity	.166	.209	.183	.168	.168	-.113	-.144	-.133
price (US$/kWh)	[.12–.194]		[.162–.209]	[.153–.184]	[.124–.191]	(.073)	(.095)	(.096)
Share of households	.584	.479	.58	.61	.543	.002	-.09*	-.041
with TV	[.477–.682]		[.45–.701]	[.511–.666]	[.436–.64]	(.053)	(.048)	(.047)
Share of households	.098	.179	.108	.043	.044	-.097	.103	.23
with computer	[.04–.155]		[.029–.234]	[.013–.07]	[.011–.078]	(.109)	(.141)	(.172)
Share of households	.071	.311	.06	.044	.125	n.a.	n.a.	n.a.
with air conditioner	[.005–.183]		[.008–.311]	[.014–.08]	[.047–.228]			
Average household	3.47	3.29	3.51	4.15	4.38	.005	-.028***	.001
size	[3.18–3.66]		[3.19–3.74]	[3.93–4.54]	[3.9–4.85]	(.006)	(.007)	(.013)
Average number of	6.22	5.47	5.81	5.57	4.53	.007	.019	.034*
rooms in dwelling	[5.73–6.71]		[5.27–6.41]	[5.02–5.77]	[4.01–5.27]	(.011)	(.012)	(.02)
Share of dwellings	.919	.974	.95	.656	.536	-.025	.222***	.341***
with bathroom	[.809–.971]		[.793–.994]	[.335–.855]	[.386–.76]	(.019)	(.067)	(.047)
Observations	17	1	26	11	8	86	56	50
Share of adults	.693	.614	.659	.579	.603	.004	-.104***	-.105**
with a job	[.633–.774]		[.614–.737]	[.55–.607]	[.567–.696]	(.015)	(.031)	(.053)
Share of adults	.306	.304	.293	.16	.149	-.016	.022	.181**
formally employed	[.251–.372]		[.17–.46]	[.091–.212]	[.068–.197]	(.041)	(.054)	(.089)
Share of adults	.129	.005	.1	.151	.131	.027	.03	.065
with agricultural job	[.041–.266]		[.002–.283]	[.051–.241]	[.03–.254]	(.077)	(.09)	(.101)
Observations	17	1	26	11	7	86	56	48

Utility-level administrative data for distribution utilities in the four subsystems in 2000 and 2010 and census data matched to the concession area of these utilities in the same years. Columns (1)–(5) display descriptive statistics in 2000 (prior to the crisis) for the variables listed in the left–hand side column for distribution utilities in the South (column 1), in the Southeast/Midwest (LIGHT in column 2, all distribution utilities in column 3), in the Northeast (column 4), and in the North (column 5). Columns (6)–(8) display estimates of a long–term difference–in–difference estimator comparing the logarithm of these variables in 2010 vs. 2000 for distribution utilities in the Southeast/Midwest (column 6), in the Northeast (column 7), and in the North (column 8) compared to distribution utilities in the South. Significance levels: *10%, **5%, ***1% (s.e. clustered by utility). Regressions include fixed effects for each distribution utility and each year. All monetary values are expressed in US$2012.