Web Appendix (for online publication only)

This Web Appendix provides all the additional details and results that are mentioned but not presented in the paper. It is long, but it is not meant to be read in its entirety and we tried to organize it in such a way that a reader could quickly find the information that she is looking for.

A Main additional results mentioned in the paper

This section presents the main additional results mentioned in the text of the paper.

- The continuation of Table 1 (mentioned in Section 3.1).
- A table mentioned in footnote 40 in which we show that the estimated long-term effect of the energy-saving program in Section 3.3 is robust to controlling for covariates that are not available at the yearly level but are available in the 2000 and 2010 censuses (median household income, population size, share of households living in urban areas, average household size, average dwelling size, share of dwellings with a bathroom, employment rate).
- A figure mentioned in Section 2.1 and footnote 41 showing that trends in the main residential electricity tariff did not evolve differentially for distribution utilities in the Southeast/Midwest compared to distribution utilities in the South after the crisis.
- A figure mentioned in Section 4.1C showing that most of the customers in the sample of Figure 5c consumed well below their quota.
- A table mentioned in Section 4.2.2 showing that our short-term results in Table 3 are robust if we extend our sample of movers by only requiring them to be observed until the end of 2002
- A table mentioned in Section 4.2.3 supporting our causal interpretation and our exclusion restriction for the results in Table 3, in which we present placebo estimates using customers who moved into their housing units in similar months but after the crisis (2002-2003, 2003–2004, and 2004-2005) rather than in 2000-2001. The average consumption of same-week movers in their first 3 billing months does not predict differences in consumption levels in these samples.
- Four tables mentioned in Section 5 complementing the qualitative evidence in Table 4: a table with the information used to calculate the inputted average monthly electricity use in 1999 in Panel A; a table completing Panel B with additional appliances; a table providing information on the self-reported change in lightbulbs in Panel B; and a table presenting reported access to information about energy efficient appliances in the Southeast/Midwest and in the South after the crisis.
- A table presenting our estimates of the price elasticity of residential electricity use, which we mention in Section 6 and use to compute some of the welfare implications in that section, with details regarding our empirical strategy.

Table A.1: Additional descriptive statistics for distribution utilities in the four subsystems

	Descriptive statistics in 2000 Mean [min-max]					Differential trends 2010 vs. 2000 Coefficient, log points (s.e.)		
	South	LIGHT	Southeast/Midwest	Northeast	North	SE/MW vs. S	NE vs. S	N vs. S
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Average residential electricity	.17	.214	.187	.172	.172	-.113	-.141	-.138
price (US$/kWh)	[.123–.198]		[.166–.214]	[.156–.189]	[.127–.196]	(.073)	(.094)	(.098)
Average temperature	18.1	21.8	21.3	25	26	-.033**	-.034**	-.02
(degrees Celsius)	[17–19.8]		[18.8–24.6]	[23.1–26.6]	[25.1–26.5]	(.017)	(.017)	(.02)
Share of households	.098	.179	.108	.043	.044	-.097	.103	.23
with computer	[.04–.155]		[.029–.234]	[.013–.07]	[.011–.078]	(.109)	(.141)	(.172)
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)
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	8	86	56	50

The table uses 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 distribution utilities in the same years (the 2010 census does not have information on air conditioners; the temperature data come 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 utility 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-differences 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 distribution utility). Regressions include fixed effects for each distribution utility and each year. All monetary values are expressed in US$ from 2012.

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

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable: Log(yearly average residential electricity use)						
Treat × Year2010	-.113***	-.117***	-.103***	-.12***	-.118***	-.117***
	(.022)	(.026)	(.029)	(.026)	(.029)	(.041)
Log main tariff (R$)		-.197***	-.143		-.211**	-.16
		(.072)	(.095)		(.09)	(.11)
Log median household income (R$)		.139	.314**		.144	.431**
		(.095)	(.129)		(.126)	(.187)
Clusters	86	86	86	70	70	70
Restricted sample	No	No	No	Yes	Yes	Yes
Other controls	No	No	Yes	No	No	Yes

Units of observation: distribution utilities (defined as of 2000) in the Southeast/Midwest and in the South. The table uses utility–level data on the average residential electricity use and on the main residential electricity tariff (ANEEL administrative data) matched to household level data from the 2000 and 2010 censuses in the concession area of each distribution utility. It displays the results from estimating variants of the difference-in-differences specification in equation (14). Significance levels: *10%, **5%, ***1% (s.e. clustered by distribution utility). Regressions include distribution utility and year fixed effects. Columns (4)-(6) show that results are essentially identical as in columns (1)-(3) when we improve the comparability of distribution utilities in the two subsystems by excluding "outliers" from the Southeast/Midwest. In particular, we exclude the 8 distribution utilities with levels of key variables in 2000 (average residential electricity use, main residential electricity tariff, median household income) falling outside the range of values observed for these variables in the South in the same year (see range of values in Table 1).

Long-term difference-in-differences results controlling for census data

As mentioned in footnote 40, we show in Table A.2 that the estimated long-term effect of the energy-saving program in Section 3.3 is robust to controlling for covariates that are not available at the yearly level but are available in the 2000 and 2010 censuses (median household income, population size, share of households living in urban areas, average household size, average dwelling size, share of dwellings with a bathroom, employment rate). We use the following specification:

$$Y_{i,t} = \alpha_i + \gamma \cdot \mathbb{1}\{t = 2010\} + \delta \cdot Treat_i \cdot \mathbb{1}\{t = 2010\} + X_{i,t}\beta + \nu_{i,t}, \quad \text{for } t = \{2000, 2010\} \tag{14}$$

The estimated long-run effect is similar when we don't include any control, when we control for the main electricity tariff and median household income, 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. Moreover, we show graphically (see Sections G to J below) that long-term changes in consumption levels are systematically lower for distribution utilities in the Southeast/Midwest than for those in the South for given baseline levels or long-term changes in all the variables in Table 1 and Table A.1.

Trends and synthetic control estimates for electricity tariffs

We show here that trends in the main residential electricity tariff did not evolve differentially for distribution utilities in the Southeast/Midwest and in the South after the crisis.

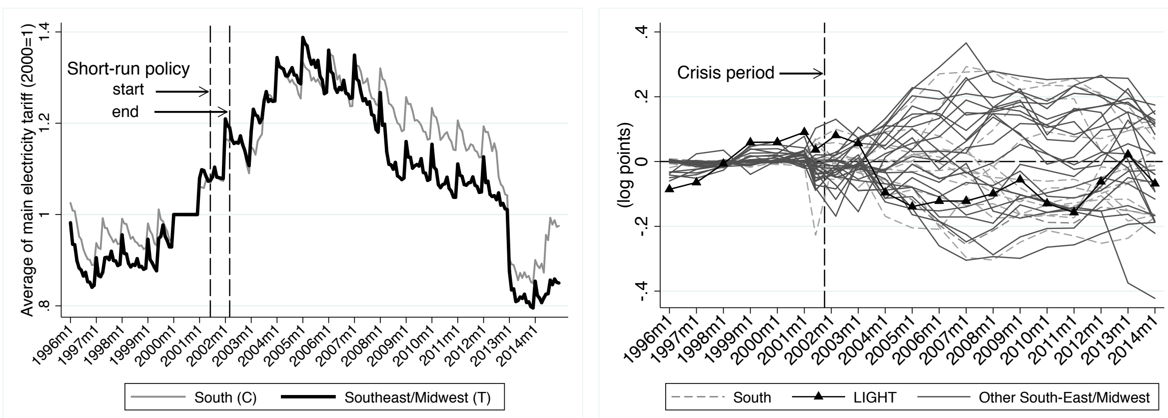
Figure A.1a is constructed similarly as Figure 1, but it displays the average of the main residential electricity tariff for distribution utilities in the Southeast/Midwest and in the South, instead of the average residential electricity use. It shows that electricity tariffs followed a similar trend in the two subsystems. If anything, electricity tariffs increased relatively less in the Southeast/Midwest (in later years). Note that we only have data on electricity tariffs starting in 1996.

Figure A.1b displays utility-specific impacts obtained by synthetic control methods, as in Figure 4b, but for the main residential electricity tariff instead of the average residential electricity use. It shows that the distributions of utility-specific impacts for the main residential electricity tariff completely overlap between distribution utilities in the Southeast/Midwest and in the South.

Figure A.1: Trends in the main residential electricity tariff

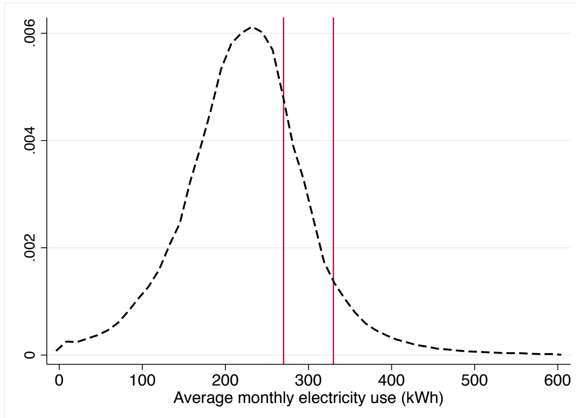
(a) Average of (real) main residential electricity tariff

(b) Synthetic control estimates for the main electricity tariff (in logs)



Panel (a) displays the (unweighted) average of the main residential electricity tariff (in real terms) for distribution utilities in the Southeast/Midwest and in the South, normalized with respect to the same month in 2000 (in order the net out any seasonality effect). Panel (b) displays utility-specific impacts obtained by synthetic control methods for the demeaned logarithm of the main residential electricity tariff. Monthly estimates are averaged into the same time periods as in Figure 4b (but starting in 1996, the first year for which we have data on electricity tariffs). 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.

Figure A.2: Consumption levels during the crisis compared to the quota for the balanced panel of customers with the same quota level (300 kWh/month) in Figure 5c



The figure uses individual monthly billing data for the universe of residential customers of LIGHT, a distribution utility subject to the energy-saving program during the crisis, from January 2000 to December 2005. It displays the consumption levels during the crisis corresponding to the "crisis" patterns in Figure 5c and show that consumption levels were well below the quota for most customers. Vertical lines show the lowest and the highest possible quota in this sample. Most customers reduced consumption below the lowest possible quota.

Complementing the evidence in Figure 5c (see Section 4.1C)

We show here that most of the customers in the sample of Figure 5c consumed well below their quota.

Figure A.2 displays the consumption levels during the crisis corresponding to the "crisis" pattern in Figure 5c and show that consumption levels were well below the quota for most customers. Vertical lines show the lowest and the highest possible quota in this sample. Most customers reduced consumption below the lowest possible quota.

Table A.3: The impact of quota on consumption – Robustness with Movers observed until 2002

	Log Quota	Log Consumption	Log Consumption July-December		
		Mar-May 2001	2001		2002
	(1)	(2)	(3)	(4)	(5)
Panel A. *Average effect*					
Log Moving-Week Quota	.811***	-.008	.13***	.136***	.114***
	(.216)	(.024)	(.034)	(.04)	(.029)
Panel B. *Median effect*					
Log Moving-Week Quota	.699***	.004	.127***	.167***	.111***
	(.011)	(.014)	(.014)	(.017)	(.016)
Log Cons. Mar-May 2001				Y	Y

The table shows the effect of quasi-experimental variation in quotas among LIGHT customers whose first bill was sent between March 2000 and February 2001 ("movers") on their electricity use in different periods. Differently from Table 3, we use the panel of movers observed regularly until December 2002 (N = 66,037) and not until December 2005. See notes in Table 3 for details. Results are comparable (a bit larger) to the ones in Table 3.

Complementing the evidence in Table 3 (see Section 4.2)

We show here some additional results complementing the evidence in Table 3.

First, Table A.3 shows that our short-term results in Table 3 are robust if we extend our sample of movers by only requiring them to be observed until the end of 2002.

Second, Table A.4 supports our causal interpretation and our exclusion restriction for the results in Table 3. It displays placebo estimates using customers who moved into their housing units in similar months but after the crisis (2002-2003, 2003–2004, and 2004-2005) rather than in 2000-2001. The average consumption of same-week movers in their first 3 billing months does not predict differences in consumption levels in these samples.

Table A.4: The impact of quota on consumption – PLACEBO using households who moved after 2002

	Log Inputted Quota	Log Consumption Mar-May	Log Consumption Jul-Dec			
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. *Movers between Mar 2002 and Feb 2003*						
		2003	2003	2003	2004	2005
Log Moving-Week Inputted Quota	1.784***	-.016	-.037	-.022	-.063	-.046
	(.074)	(.022)	(.036)	(.044)	(.05)	(.061)
Panel B. *Movers between Mar 2003 and Feb 2004*						
		2004	2004	2004	2005	
Log Moving-Week Inputted Quota	2.065***	-.022	-.076	-.056	-.074	
	(.119)	(.027)	(.047)	(.043)	(.056)	
Panel C. *Movers between Mar 2004 and Feb 2005*						
		2005	2005	2005		
Log Moving-Week Inputted Quota	1.697***	.015	-.02	-.034		
	(.146)	(.053)	(.069)	(.054)		
Log Cons. Mar-May baseline				Y	Y	Y

This table shows the results of *placebo* correlations with the average quota of same-week movers. It uses households who moved into a new housing unit after the end of the energy-saving program, as indicated in each panel. In these samples, we input a quota to each household by replicating the quota assignment rule as if each household had moved in the same week (e.g., first week of January) but in the year just prior to the crisis: in the year prior to the crisis, the assignment rule generated variation in quotas entirely due to different moving dates because of seasonality in electricity use (see Table 3 and text for details). The table then presents the result from estimating the same specifications as in Table 3 but using the (inputted) quota of mover i and the average (inputted) quota of all movers (excluding i) who received their first bill in the same week as i. *Panel A* presents estimates for a placebo crisis starting in June 2003, using the sample of Light customers whose first bill was sent between March 2002 and February 2003 (N = 41,249). *Panel B* presents estimates of a placebo crisis starting in June 2004, using the sample of Light customers whose first bill was sent between March 2003 and February 2004 (N = 43,333). *Panel C* presents estimates of a placebo crisis starting in June 2005, using the sample of Light customers whose first bill was sent between March 2004 and February 2005 (N = 10,704). In all three panels we restrict attention to movers whose average consumption in the three months prior to the placebo crisis was in the top quartile of the distribution, and whom we observe regularly until the end of 2005. Significance levels: *10%, **5%, ***1%. In all placebo exercises, the average (inputted) quota of same-week movers predict the movers' own (inputted) quota in the placebo crisis (column 1) but not the subsequent consumption levels (columns 2-6). We find no effect on average consumption levels prior (column 2), during (columns 3 and 4) or after (columns 5 and 6) the placebo crisis.

Complementing the evidence in Table 4 using PPH surveys (see Section 5)

We provide here some complementary evidence and information for Table 4 using PPH surveys (see Section 5).

First, Table A.5 displays the average electricity use of appliances in 1999 in the Southeast/Midwest for hypothetical utilization rates, as calculated by the Electrical Energy Saving Program of the Brazilian government (PROCEL). Second, Table A.6 is similar to Panel B in Table 4, but it presents reported changes in the utilization of "Other" and "Stand By" appliances for households in the Southeast/Midwest after the crisis. Third, Table A.7 presents the share of households in the Southeast/Midwest that reported adopting CFL lightbulbs during the crisis and that kept using them afterward. Finally, Table A.8 presents reported access to information about energy efficient appliances in the Southeast/Midwest and in the South in 2005. The differences between the two groups are small. Households in the Southeast/Midwest were more likely to report that they received information about energy efficiency, but were actually less likely to know what key information for the energy efficiency of electrical appliances in Brazil meant or how it could be used. Unfortunately, the 1999 survey did not ask these questions.

Table A.5: Inputted Average Electricity Consumption by Electrical Appliance

	Appliance Specification	Hypothetical utilization rate (1)	Average Monthly Consumption (kWh) (2)
Electric Shower	Low Power	18 minutes/person/day*	61.0
	High Power	18 minutes/person/day*	87.1
Refrigerator	1 Door, Frost Free	24 hours/day	42.8
Freezer		24 hours/day	43.5
Lightbulbs	Incandescent 60 Watts	5 hours/day	10.2
	Fluorescent 15 Watts	5 hours/day	2.25
TV		5 hours/day	13.5
Air Conditioner	Wall, 9001-14000 BTU	24 hours/week	69.0
Washing M.		12 loads/month	30.9
Microwave		20 minutes/day	14.0

This table presents the average electricity use of appliances for hypothetical utilization rates, as calculated by the Electrical Energy Saving Program of the Brazilian government (PROCEL). These figures are based on technical characteristics of appliances and hypothetical utilization rates drawn from the PPH survey.
(*) The shower calculation is based on a household size of 3.25 household members using the shower (number obtained from PPH 1998-1999). The complete table can be found at the website www.eletrobras.com/procel.

Table A.6: Reported mechanisms of long-run changes in electricity use

	Other Appliances (1)	Stand By Appliances (2)
Conditional on owning a given type of appliance before the crisis, share of respondents who report that they:		
(1) Use appliance as much as before crisis	.24	.68
(2) Use appliance less than before crisis	.71	.27
(3) Disconnected or disposed of appliance	.03	.05
(4) Substituted a more energy-efficient model	.02	0
Number of respondents	63	3,325

The table uses household-level data for 8 distribution utilities in the Southeast/Midwest from the most recent rounds of the PPH surveys (2004/2005). It tabulates the share of households owning "Other appliances" (besides those considered in Panel B in Table 4) and "Stand By" appliances prior to the crisis that chose each statement as the best answer to a question about their usage of the appliances after the crisis.

Table A.7: Adoption of more efficient lightbulbs around the crisis

	All (1)	Some (2)	None (3)	Number of respondents
Did you substitute incandescent lightbulbs with fluorescent ones?	.29	.14	.51	4,648
Do you still use fluorescent lightbulbs?	.60	.09	.26	1,963

The table uses household-level data for 8 distribution utilities in the Southeast/Midwest from the most recent rounds of the PPH surveys (2004/2005). It tabulates the share of households choosing each answer (in columns) for two questions about their adoption of fluorescent lightbulbs (in rows).

Table A.8: Access to information about the energy efficiency of electrical appliances

	Mean		Difference
	Southeast/Midwest	South	= (1) - (2)
	(1)	(2)	(3)
Do you receive information about energy-efficient appliances and energy-saving measures?	.73	.62	.11
			(.33)
Do you know the label for energy-efficient appliances (PROCEL)?	.46	.50	-.04
			(.11)
Do you know what the PROCEL label represents?	.34	.40	-.06
			(.14)
Do you know how much you can save by using labeled appliances?	.21	.21	0
			(.01)

The table uses household-level data for 10 distribution utilities in the Southeast/Midwest (8) and in the South (2) from the most recent rounds of the PPH surveys (2004/2005). It reports the share of households who answered "yes" to different questions about access to information (in rows) in the Southeast/Midwest (column 1) and in the South (column 2). Column 3 presents the estimated difference between these two columns. Significance levels: *10%, **5%, ***1% (s.e. clustered by distribution utility and estimated using the wild-cluster bootstrap-t). N=3,364.

Estimating the price elasticity of residential electricity use

In Section 6 in the paper, we use an estimate of the price elasticity of average residential electricity use in the Southeast/Midwest. We detail here how we obtain this estimate.

We use the utility-level panel data from ANEEL for the distribution utilities in the Southeast/Midwest. Demand typically responds with a lag, so we average all variables at the yearly level. We are interested in a price elasticity after the crisis, so we only consider data from 2003 onward.[58] We regress the logarithm of average residential electricity use $y_{i,t}$ on the logarithm of the main residential tariff $tariff_{i,t}$:

$$y_{i,t} = a_i + \beta_t + \eta \cdot tariff_{i,t} + X_{i,t}\gamma + \nu_{i,t} \tag{15}$$

where a_i and β_t are fixed effects for distribution utility i and year t. $\nu_{i,t}$ is an error term clustered by utility. $X_{i,t}$ are yearly controls for total population, total formal employment, GDP, and average temperature for each distribution utility (in log; we use data until 2012 given the availability of such controls). η captures our price elasticity.

There are two major concerns with an equation such as equation (15). First, there is rarely a unique price of electricity. In Brazil, the main residential tariff is essentially linear, but an alternative tariff for low-income and small consumers offers nonlinear percentage discounts on this unit price. Changes in residential prices, however, typically apply to the main tariff. Therefore, percentage changes in the main tariff capture percentage changes in every marginal price (thus the log specification).

Second, changes in prices may be endogenous to changes in quantities. The price-cap mechanism limits such a concern in Brazil. Between *revision* years, demand risk falls entirely on distribution utilities and, by design, yearly price *adjustments* are not endogenous to changes in quantities (ANEEL, 2005). Price revisions every 4 to 5 years may still create some endogeneity, biasing estimates of η away from 0. We directly assess the extent of endogeneity in two ways. First, we run the same regression instrumenting the main tariff by its cost-of-energy component (exogenous to the firm on a yearly basis), available for every utility since 2005. Second, we estimate equation (15), excluding years of price *revisions* and including utility-specific fixed effects for each between-revision period. The only variation left comes from price *adjustments*.

Results are presented in Table A.9. We estimate $\widehat{\eta}$ at $-.2051$ (column 1) and $-.1942$ (column 4) with the full variation in tariffs from 2003 and 2005, respectively. Estimates using only the variation from price adjustments (column 3) are similar (because sample years are different in column 3, we also show results from a similar specification as in column 1 for those sample years

[58] In practice, there is not much price variation prior to the crisis because the electricity sector was liberalized in the second half of the 1990s and prices did not start to vary much in the first few years after the liberalization.

Table A.9: Price elasticity estimates using yearly variation in the South-East/Midwest post-crisis

Dependent variable: Log(yearly mean of average residential electricity use)					
	(1)	(2)	(3)	(4)	(5)
Log(yearly mean of main residential tariff)	-.2051***	-.2094***	-.2232***	-.1942***	-.3099**
	(.03782)	(.03632)	(.06335)	(.04991)	(.153)
First stage dependent variable: Log(yearly mean of main residential tariff)					
	(1)	(2)	(3)	(4)	(5)
Log(yearly mean of the cost of energy in the main residential tariff)					.2508***
					(.06314)
Model	OLS	OLS	OLS	OLS	IV-2SLS
First year	2003	2003	2003	2005	2005
Exclude variation from revision years	No	Yes	Yes	No	No
Between-revision FE	No	No	Yes	No	No
Observations	270	201	201	216	216
Clusters	27	27	27	27	27

This table presents the price elasticity estimates used in Section 6. Monthly observations for every distribution utility in the Southeast/Midwest are averaged out by year. The sample includes observations after the crisis, from 2003 until 2012. All regressions control for year and distribution utility fixed effects, as well as population, formal employment, GDP per capita, and average temperature (yearly, in log) for each distribution utility. Significance levels: *10%, **5%, ***1% (s.e. clustered by distribution utility).

in column 2). The estimated elasticity is higher with the IV strategy (column 5), at $-.3099$ (the first stage is strong). In the paper, we use estimates from column (5), which is a conservative choice for our computation of the welfare implications of hysteresis: the welfare effects and the bias from assuming away hysteresis would be larger with a smaller elasticity.

B Extended Conceptual Framework

This section presents an extended version of Section 1 containing lengthier discussions of our assumptions and our results, as well as all our derivations.

We begin by showing how hysteresis affects the welfare evaluation of corrective policies. Hysteresis is defined again as a situation in which a change in a behavior – we use electricity use as running example – in a given period persists in subsequent periods, even though the force inducing the initial change no longer applies; in other words, when levels of a behavior are complements across time periods. Our approach in this section is to remain as close as possible to the textbook price-theory framework that has been used to derive canonical results in the economics of corrective policies (e.g., optimal Pigouvian taxation). This allows us to highlight the welfare implications of hysteresis in a familiar setting. Moreover, the results we derive using this framework have the advantage that they generalize beyond a specific model, as long as slopes of demand (and supply) curves for the behavior subject to a corrective policy are *sufficient statistics* for changes in private welfare (i.e., total surplus) due to changes in that behavior. In fact, we also extend our analysis to models for which these slopes may fail to be sufficient statistics.

We start by laying out a benchmark framework and by discussing mechanisms of hysteresis it applies to. We then derive sufficient-statistics formulas for the welfare effect of corrective policies. In so doing, we show how estimates of the persistent effect of short-run policies can inform the welfare effect of short-run (i.e., temporary) and long-run (i.e., permanent) policies, whether the government has a short-run goal (e.g., curbing energy demand in the face of a temporary shortage) or a long-run goal (e.g., reducing energy demand to mitigate climate change). Next, we consider extensions to the benchmark framework. In each case, we show that we obtain canonical formulas for the welfare effect of corrective policies when assuming away hysteresis. We then obtain new formulas allowing for hysteresis, and we highlight the bias from failing to take it into account.

Although we use household electricity use as running example throughout the section, note that the theory applies equally well to other household behaviors for which the possibility of hysteresis may be relevant. Moreover, our results are not specific to settings in which households are on the demand side of the market.[59] Finally, our results are not specific to inter-temporal complementarities, but carry over to policies correcting behaviors that are complements within a given time period, rather than across time periods as for hysteresis (see footnotes throughout the section).

[59] E.g., for an application on charitable donation, we would model the household on the supply side of the market and we would obtain the same results (appropriately replacing references to "demand curves" by "supply curves").

B.1 General setup and mechanisms

We make some assumptions to focus on the implications of hysteresis for the welfare evaluation of corrective policies. First, we consider the problem of a representative household deciding how much electricity to use in each of two periods. A two-period setting is sufficient for our purpose, and we allow for heterogenous households in an extension. Second, we assume that electricity use generates externalities in order to motivate the need for corrective policies. The government has a short-run (resp. long-run) goal when the externality applies to the first period only (resp. to both periods). Third, we adopt a partial-equilibrium setup to focus on the household decision, and we use quasi-linear utility functions to assume away income effects and redistributive concerns, as is standard in the literature (e.g., Allcott and Taubinsky, 2015). Fourth, we assume that electricity is priced at a constant marginal cost, such that we can illustrate the role of hysteresis using simple consumer surplus concepts. We introduce a producer surplus in an extension. Fifth, we model the household as fully rational and forward looking; doing so is necessary for slopes of the demand curves to be sufficient statistics for changes in consumer surplus. We later relax these assumptions.

We introduce hysteresis in this rather standard setup by allowing changes in electricity use due to a corrective policy in period 1 to affect the household's propensity to use electricity in period 2, or the energy (in-)efficiency of the mapping from its electricity use to the utility it derives from it.

As we show below, it is useful to distinguish between two categories of mechanisms of hysteresis. First, changes in electricity use in period 1 may result from *active* investments that modify the household's propensity to use electricity persistently. In our context, this category encompasses physical investments, e.g., in home improvements or in the energy efficiency of the stock of electrical appliances. It also encompasses investments in more intangible assets, such as information acquisition, e.g., learning about ways to use electricity more efficiently for a given appliance stock or about the pros and cons of new technologies. In both cases, the household's propensity to use electricity may change persistently because investment costs are not easily reverted (e.g., the resale value of appliances drops quickly and learning costs are typically sunk) and the assets created do not depreciate immediately. Second, changes in electricity use in period 1 may have long-run effects on the household's propensity to use electricity because of its history of consumption. Like in experience-good models, the household may learn about the costs and benefits of adopting new behaviors by experimenting with different levels of electricity use in period 1, acquiring new information, or developing new habits *passively* (e.g., Bryan, Chowdhury and Mobarak, 2014; Dupas, 2014). Habit formation models, in which the marginal utility of consumption depends on past consumption levels, belong to this category (Becker and Murphy, 1988).[60]

Our framework captures these two categories in a "reduced-form" fashion such that our re-

[60]The optimal consumption path in rational habit formation models may have multiple steady states, and a corrective policy that pushes the household far enough from its prior steady state may lead to hysteresis.

sults are not tied to modeling a particular mechanism. In fact, our results generalize to any such mechanism, as long as slopes of the demand curves remain sufficient statistics for changes in the household welfare from having to deviate from its privately optimal choices. This condition is a natural starting point to focus on the implications of hysteresis, as it is the same condition used to derive the canonical results for optimal Pigouvian taxation. We show that, under this condition, we do not need to know which specific mechanisms cause the persistent effect of a short-run policy for the welfare evaluation of corrective policies, as long as we can quantify the degree of hysteresis.[61]

B.2 Formal presentation of the benchmark framework

We now present the benchmark framework formally. The household derives utility from the services provided by electricity use x_t, with unit price p_t, and from a numeraire c_t, given its income y_t in periods $t = 1, 2$. The per-period utility from electricity services is $v_t(x_t, s_t)$ and is strictly increasing and concave in x_t. This function captures the mapping from electricity use to utility, i.e., electricity use provides services that create utility. It implicitly involves usage and electrical appliance choices, e.g., consuming more electricity implies adding appliances or new ways to use these appliances.[62] s_t is the household's "propensity to consume" electricity. It captures the energy (in-)efficiency of the mapping from electricity use to utility, i.e., usage and appliance choices can be more or less energy-efficient. Without loss of generality, high values of s_t imply a high propensity to consume (e.g., energy-inefficient appliances or habits, limited knowledge of energy-efficient behaviors). We make two key assumptions about this variable that are also made in, e.g., Becker and Murphy (1988). First, we assume $\frac{\partial v_t}{\partial s_t} \leq 0$: the household derives less utility from a given level of electricity use if its stock of appliances, its habits, or its knowledge are less energy-efficient (this assumption essentially defines s_t). Second, we assume complementarity, $\frac{\partial^2 v_t}{\partial x_t \partial s_t} > 0$: the higher the propensity to consume (e.g., the more electricity the household needs to obtain services), the higher the marginal utility from electricity use (e.g., the greater the benefits from more electricity).

We introduce hysteresis by allowing the propensity to consume to be a function $s_t(s_{t-1}, x_{t-1}, I_t)$. The household can make investments, I_t, to reduce its concurrent propensity to consume, $\frac{\partial s_t}{\partial I_t} \leq 0$, at strictly increasing and convex costs $\kappa_t(I_t)$, which may or may not be monetary (e.g., price of new appliances or time devoted to learning).[63] These investments can affect the future propensity

[61] We implicitly assume that there are no (differential) pre-existing distortions on the household's various margins of adjustment. Otherwise, it would be necessary to know which mechanisms cause hysteresis. For instance, if the manufacturing of appliances generates its own negative externality (resp. appliances are taxed), a change in electricity use coming from the purchase of new appliances will be more distortionary (resp. less distortionary) than if it was coming from the adoption of new behaviors because it will worsen (resp. mitigate) a pre-existing distortion. Yet, this qualifier applies even in absence of hysteresis. Moreover, it is straightforward that our main insights would survive; we would simply have to quantify these additional distortions and the persistent effect attributed to each mechanism.

[62] It is common in the literature to not specify all the choices involved in such a mapping (see, e.g., Ito, 2015).

[63] The two representations are equivalent with quasi-linear utility. Our model is consistent with Dubin and McFadden

to consume because we assume $\frac{\partial s_t}{\partial s_{t-1}} \geq 0$ (this allows for depreciation). Consumption choices can also affect the household's future propensity to consume directly, $\frac{\partial s_t}{\partial x_{t-1}} \geq 0$. The costs of reducing its propensity to consume through this channel are already captured by assuming $\frac{\partial v_t}{\partial x_t} > 0$. Changes in the future propensity to consume lead to hysteresis because of the complementarity assumption.

The household then solves the following problem:

$$\max_{x_1,x_2,I_1,I_2} V = U_1 + \beta U_2 = y_1 - p_1 x_1 - \kappa_1(I_1) + v_1(x_1,s_1) + \beta\left[y_2 - p_2 x_2 - \kappa_2(I_2) + v_2(x_2,s_2)\right] \tag{16}$$

$$\text{s.t. } s_t = s_t(s_{t-1}, x_{t-1}, I_t) \text{ for } t = 1,2 \text{ and } \{s_0, x_0\} \text{ given,}$$

where V is the household "lifetime" utility and β accounts for discounting and possible differences in the relative length of the two periods. The first-order conditions for this problem are:

$$\frac{\partial v_1(x_1,s_1)}{\partial x_1} + \beta \frac{\partial v_2(x_2,s_2)}{\partial s_2}\frac{\partial s_2}{\partial x_1} = p_1 \quad ; \quad \frac{\partial v_2(x_2,s_2)}{\partial x_2} = p_2 \tag{17}$$

$$\left[\frac{\partial v_1(x_1,s_1)}{\partial s_1} + \beta \frac{\partial v_2(x_2,s_2)}{\partial s_2}\frac{\partial s_2}{\partial s_1}\right]\frac{\partial s_1}{\partial I_1} - \kappa_1'(I_1) = 0 \quad ; \quad \frac{\partial v_2(x_2,s_2)}{\partial s_2}\frac{\partial s_2}{\partial I_2} - \kappa_2'(I_2) = 0 \tag{18}$$

The first-order conditions highlight the costs and benefits of marginal changes in electricity use and investments, respectively. They imply that the household uses more electricity if its concurrent propensity to consume is higher ($\frac{\partial^2 v_t}{\partial x_t \partial s_t} > 0$); that it uses less electricity in period 1 if its future propensity to consume depends on past choices ($\frac{\partial s_2}{\partial x_1} > 0$); and that it invests more in period 1 if investments are persistent ($\frac{\partial s_1}{\partial I_1} < 0$ and $\frac{\partial s_2}{\partial s_1} > 0$). These first-order conditions and baseline electricity prices (p_{10} and p_{20}) determine baseline electricity use and investment levels (x_{10}, x_{20}, I_{10}, and I_{20}).

B.3 Social welfare and corrective policies

We motivate the need for corrective policies in this framework by assuming that social welfare, W, differs from private welfare or the household's lifetime utility, V. In particular, there are externalities from electricity use, $E_t(x_t)$, which we assume to be negative without loss of generality, such that welfare can be written $W = V - E_1(x_1) - \beta E_2(x_2)$, with $E_t(x_t)$ positive and increasing in x_t.

We then evaluate the welfare effect of two types of policies. First, we consider a short-run policy, in which the government restricts electricity use to $\overline{x_1} < x_{10}$ in period 1 only. The welfare

(1984), where investing in a more energy-efficient technology increases the marginal productivity of electricity. Our assumption that this decreases the marginal utility of electricity use ($\frac{\partial^2 v_t}{\partial x_t \partial s_t}\frac{\partial s_t}{\partial I_t} \leq 0$) is possible in Dubin and McFadden (1984), as is the opposite. We make our assumption to introduce the possibility of hysteresis in the direction that is consistent with our empirical evidence, i.e., that a policy inducing the household to reduce electricity use in period 1 may induce the household to reduce electricity use in period 2. However, our sufficient statistics formulas do not rely on this assumption: they would still hold if the persistent effect of the short-run policy was instead to increase electricity use, as long as we can estimate this effect. That we model investments as a continuous variable is inconsequential; we only need the aggregate demand curve for electricity to be continuous.

effect of this policy is $\Delta W^{SR} = W\left(\overline{x_1}, x_2(\overline{x_1}), I_1(\overline{x_1}), I_2(\overline{x_1})\right) - W\left(x_{10}, x_{20}, I_{10}, I_{20}\right)$, where SR stands for "short run." We denote by $x_2(\overline{x_1})$, $I_1(\overline{x_1})$, and $I_2(\overline{x_1})$, privately optimal levels of these variables given the short-run policy and the price p_{20}. Second, we consider a long-run policy, in which the government restricts electricity use to $\overline{x_1} < x_{10}$ in period 1 and to $\overline{x_2} < x_{20}$ in period 2. The welfare effect of this policy is $\Delta W^{LR} = W\left(\overline{x_1}, \overline{x_2}, I_1(\overline{x_1}, \overline{x_2}), I_2(\overline{x_1}, \overline{x_2})\right) - W\left(x_{10}, x_{20}, I_{10}, I_{20}\right)$, where LR stands for "long run." We denote by $I_1(\overline{x_1}, \overline{x_2})$ and $I_2(\overline{x_1}, \overline{x_2})$, privately optimal levels of these variables given the long-run policy. Finally, we assume that these restrictions on electricity use are implemented through traditional instruments, i.e., through quotas or taxes (we consider a policy that includes social incentives in an extension). The welfare gain of both policies then arises from the correction of the externalities and the welfare loss from the reduction in the household private welfare from having to deviate from its privately optimal choices regarding electricity use.[64]

A short-run policy could be motivated by cases in which the slope of the externality function – the marginal damage $MD_t(x_t)$ – is larger in period 1 given baseline choices, as is the case with a temporary electricity supply crisis and the higher risk of blackouts. A long-run policy could be motivated by the need to change behaviors persistently, for instance to tackle climate change. Yet, our framework allows the government to have short-run or long-run goals when evaluating the welfare effect of these policies. Specifically, it has a short-run (resp. long-run) goal if the marginal damage is assumed to be nil in period 2 (resp. to be positive in both periods).

B.4 Welfare effect assuming away hysteresis

We now show how a researcher *who assumes away hysteresis* would evaluate the welfare effect of short-run and long-run corrective policies using a standard price-theory approach. In this case, the household's propensity to consume is assumed to not depend on past consumption choices, and investments fully depreciate from one period to the next; that is, we have $s_t\left(s_0, I_t\right)$ for a given s_0.

B.4.1 Short-run policy

The researcher can recover the welfare effect of the short-run policy by tracing it along the path from x_{10} to $\overline{x_1}$, letting I_1 adjust endogenously, as variables in period 2 are assumed to be unaffected:

$$\begin{aligned}\Delta W^{SR}_{NoH} &= \int_{x_{10}}^{\overline{x_1}} \frac{dW(x_1, x_2, I_1, I_2)}{dx_1} dx_1 = \int_{x_{10}}^{\overline{x_1}} \frac{dU_1(x_1, I_1)}{dx_1} dx_1 - \int_{x_{10}}^{\overline{x_1}} \frac{dE_1(x_1)}{dx_1} dx_1 \\ &= \int_{x_{10}}^{\overline{x_1}} \left[p_{1,NoH}(x_1) - p_{10}\right] dx_1 - \int_{x_{10}}^{\overline{x_1}} MD_1(x_1) dx_1 \end{aligned} \tag{19}$$

[64] In the case of behaviors that are complements within a given period, the short-run policy corresponds to a policy correcting one behavior only and the long-run policy to a policy correcting both behaviors at the same time.

where *NoH* stands for "no hysteresis." Equation (19) is the textbook formula for the welfare effect of a corrective policy. The welfare gain from correcting the externality corresponds to the area under the marginal damage function $MD_1(x_1)$ between the consumption levels with and without the policy in period 1. The welfare loss from reducing electricity use below the privately optimal level is captured by a Harberger triangle. It corresponds to the area under the demand curve $p_{1,NoH}(x_1)$ and above the baseline price p_{10} between the consumption levels with and without the policy in period 1, or the triangle $A_1B_1C_1$ in Figure B.1a for the case of a linear demand curve.

The key insight of price theory is that the slope of the demand curve is a sufficient statistic for the change in the household private welfare (i.e., the consumer surplus) because it accounts for changes in its utility due to changes in x_1 and to changes in other choices variables following changes in x_1 (e.g., I_1). As a result, it does not matter *how* the household reduced its electricity use. The household optimally chooses among all its margins of adjustments, and any loss in its utility is reflected in the demand curve. Therefore, if p_{10} and x_{10} are known, two empirical objects are sufficient for evaluating the welfare effect of the short-run policy. The slope of the demand curve $p_{1,NoH}(x_1)$ can be estimated using exogenous changes in electricity prices in period 1. The marginal damage function $MD_1(x_1)$ can be estimated using the various approaches in the literature.

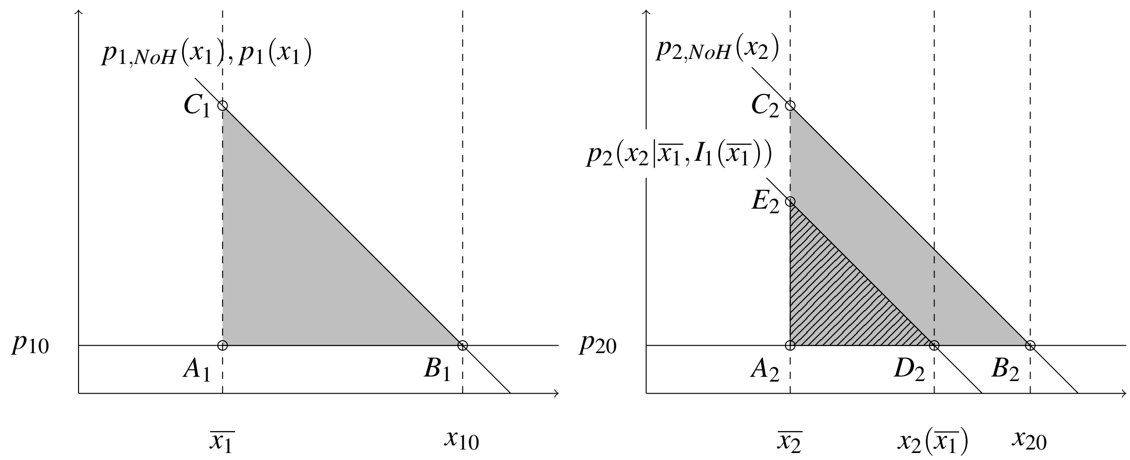
B.4.2 Long-run policy

Since choices in the two periods are assumed to be independent, the researcher can recover the welfare effect of the long-run policy by separately tracing it along the paths from x_{10} to $\overline{x_1}$ and from x_{20} to $\overline{x_2}$, letting I_1 and I_2 adjust endogenously:

$$\begin{aligned}\Delta W_{NoH}^{LR} &= \int_{x_{10}}^{\overline{x_1}} \frac{dW(x_1,x_2,I_1,I_2)}{dx_1}dx_1 + \int_{x_{20}}^{\overline{x_2}} \frac{dW(x_1,x_2,I_1,I_2)}{dx_2}dx_2 \\ &= \int_{x_{10}}^{\overline{x_1}} \frac{dU_1(x_1,I_1)}{dx_1}dx_1 + \beta\int_{x_{20}}^{\overline{x_2}} \frac{dU_2(x_2,I_2)}{dx_2}dx_2 - \int_{x_{10}}^{\overline{x_1}} \frac{dE_1(x_1)}{dx_1}dx_1 - \beta\int_{x_{20}}^{\overline{x_2}} \frac{dE_2(x_2)}{dx_2}dx_2 \\ &= \int_{x_{10}}^{\overline{x_1}} [p_{1,NoH}(x_1) - p_{10}]\,dx_1 + \beta\int_{x_{20}}^{\overline{x_2}} [p_{2,NoH}(x_2) - p_{20}]\,dx_2 - \int_{x_{10}}^{\overline{x_1}} MD_1(x_1)dx_1 - \beta\int_{x_{20}}^{\overline{x_2}} MD_2(x_2)dx_2\end{aligned} \tag{20}$$

Equation (20) is a two-period version of the textbook formula in equation (19). The total loss in private welfare corresponds to the (discounted) sum of Harberger triangles in both periods, or the areas $A_1B_1C_1$ and $A_2B_2C_2$ in Figures B.1a and B.1b, respectively. If p_{20} and x_{20} are also known, equation (20) shows that three additional objects are needed to evaluate the welfare effect of the long-run policy: the discount rate β, the marginal damage function $MD_2(x_2)$, and the slope of the demand curve $p_{2,NoH}(x_2)$, which could be estimated using exogenous changes in electricity prices once in period 2. Finally, it is straightforward from equations (19) and (20) that a government with a short-run (resp. long-run) goal will want to implement a short-run (resp. long-run) policy.

Figure B.1: The loss in private welfare from corrective policies with and without hysteresis



(a) Period 1 (b) Period 2

The figure illustrates the loss in private welfare in the benchmark framework (the private welfare corresponds to the consumer surplus) for the short-run and the long-run corrective polices. For simplicity, we assume linear demand curves and no change in the slope of the demand curve following a short-run policy in the presence of hysteresis. The triangles $A_1B_1C_1$ illustrates the loss in private welfare for the short-run policy with and without hysteresis. The demand curve $p_2\left(x_2|\overline{x_1}, I_1\left(\overline{x_1}\right)\right)$ illustrates that the demand curve shifts inward in period 2 following a short-run policy in the presence of hysteresis. The triangles $A_1B_1C_1$ and $A_2B_2C_2$ illustrate the loss in private welfare for the long-run policy assuming away hysteresis. The triangles $A_1B_1C_1$ and $A_2D_2E_2$ illustrate the comparable loss in private welfare allowing for hysteresis. The trapezoid $D_2B_2C_2E_2$ illustrates the first source of bias from assuming away hysteresis for the long-run policy. See text for more details.

B.5 Welfare effect allowing for hysteresis

We now show how a researcher *who allows for hysteresis* would evaluate the welfare effect of the two policies using price theory. In this case, the propensity to consume takes the form $s_t\left(s_{t-1}, x_{t-1}, I_t\right)$.

B.5.1 Short-run policy

The researcher can still recover the welfare effect of the short-run policy by tracing it along the path from x_{10} to $\overline{x_1}$. However, one must now recognize that x_2 and I_2 may also adjust endogenously:[65]

$$\begin{aligned}\Delta W^{SR} &= \int_{x_{10}}^{\overline{x_1}} \frac{dW(x_1,x_2,I_1,I_2)}{dx_1}dx_1 = \int_{x_{10}}^{\overline{x_1}} \frac{dV(x_1,x_2,I_1,I_2)}{dx_1}dx_1 - \int_{x_{10}}^{\overline{x_1}} \frac{dE_1(x_1)}{dx_1}dx_1 - \beta\int_{x_{10}}^{\overline{x_1}} \frac{dE_2(x_2)}{dx_2}\frac{dx_2}{dx_1}dx_1 \\ &= \int_{x_{10}}^{\overline{x_1}} \left[p_1(x_1) - p_{10}\right]dx_1 - \int_{x_{10}}^{\overline{x_1}} MD_1(x_1)dx_1 - \beta\int_{x_{20}}^{x_2(\overline{x_1})} MD_2(x_2)dx_2 \qquad (21)\end{aligned}$$

The same insight of price theory implies that the slope of the demand curve in period 1 remains a sufficient statistic for the change in private welfare due to the short-run policy, including from changes in x_2 and I_2. Moreover, the slope of the demand curve $p_1(x_1)$ could be estimated in the same way as a researcher would estimate the slope of the demand curve $p_{1,NoH}$ when assuming away hysteresis. The only novelty in equation (21) is that the researcher must now consider a

[65] Ito, Ida and Tanaka (2017) find persistent effects for 3 months after a policy ended, and use a formula similar to the one derived here to evaluate the welfare implications of that specific short-run policy (see their Web Appendix).

potential effect on the externality in period 2 because of hysteresis in electricity use. As Figure B.1b illustrates, the demand curve in period 2, $p_2(x_2|\overline{x_1}, I_1(\overline{x_1}))$, may have shifted inward given the change in the household's propensity to consume due to its choices in period 1. This shift leads to a privately optimal level of electricity use in period 2, $x_2(\overline{x_1})$, that is lower at price p_{20}.

The bias from assuming away hysteresis when evaluating the welfare effect of a short-run policy therefore comes from neglecting the possible correction of the externality in period 2:

$$Bias^{SR} = \Delta W^{SR}_{NoH} - \Delta W^{SR} = \beta \int_{x_{20}}^{x_2(\overline{x_1})} MD_2(x_2)dx_2 \leq 0 \tag{22}$$

One would underestimate the welfare gain of the policy, unless the government has only a short-run goal ($MD_2(x_2) = 0, \forall x_2 \leq x_{20}$). The magnitude of the bias depends on the marginal damage function in the long run $MD_2(x_2)$, and on the long-run effect of the short-run policy $x_{20} - x_2(\overline{x_1})$.[66]

The long-run effect of the policy is thus the new key empirical object to estimate. More precisely, it corresponds to the long-run effect of a policy that was *expected to be short run*. This qualifier is important because studies that estimate the persistent effect of a short-run policy rarely specify whether the policy was expected to last longer than it did. The government may have planned for a long-run policy, but later revised its plans. The household may have formed beliefs that the policy would persist, even if the government planned to implement the policy temporarily. Policy expectations matter because the household may make investments in period 1 (e.g., replacing an appliance or insulating the house) that it would not make if it knew that the policy were to be only temporary. These "extra" investments, which may contribute to the persistent effect of the policy in period 2, are ex-post suboptimal for the household and imply losses in private welfare that are not accounted for by the demand curve in period 1.

Formally, we can illustrate this point by considering a policy in which the government first announces electricity use restrictions $\overline{x_1} < x_{10}$ and $\overline{x_2} < x_{20}$, but then cancels the policy at the start of period 2. Let $x_2(\overline{x_1}, I_1(\overline{x_1}, \overline{x_2}))$ and $I_2(\overline{x_1}, I_1(\overline{x_1}, \overline{x_2}))$ be the household's choices in period 2 given its choices in period 1 when planning for a long-run policy, $\overline{x_1}$ and $I_1(\overline{x_1}, \overline{x_2})$. We can recover the welfare effect of such a policy that was "not expected to be short run" (*NotExp*) by tracing it along the path from x_{10} to $\overline{x_1}$. However, we must then also account for the extra investments that the

[66] This result carries over to behaviors that are complement within a time period. Suppose, for instance, that smoking and drinking generate negative externalities and that the household likes to smoke when drinking. A policy that reduces drinking would thus also reduce smoking. The demand curve for drinking captures the associated utility from smoking and so the slope of the demand curve for drinking is sufficient to measure any change in utility from the associated change in smoking. The bias from assuming this complementarity away when evaluating the welfare effect of a policy restricting only drinking therefore comes from neglecting the possible correction of externalities from smoking.

household may make in period 1 in anticipation of the period-2 policy:

$$\begin{aligned}\Delta W^{SR}_{NotExp} &= \int_{x_{10}}^{\overline{x_1}} \frac{dW(x_1,x_2,I_1,I_2)}{dx_1} dx_1 + \int_{I_1(\overline{x_1})}^{I_1(\overline{x_1},\overline{x_2})} \frac{dW(\overline{x_1},x_2,I_1 I_2)}{dI_1} dI_1 \\ &= \int_{x_{10}}^{\overline{x_1}} [p_1(x_1) - p_{10}]\, dx_1 + \int_{I_1(\overline{x_1})}^{I_1(\overline{x_1},\overline{x_2})} \frac{dV(\overline{x_1},x_2,I_1 I_2)}{dI_1} dI_1 - \int_{x_{10}}^{\overline{x_1}} MD_1(x_1) dx_1 - \beta \int_{x_{20}}^{x_2(\overline{x_1},I_1(\overline{x_1},\overline{x_2}))} MD_2(x_2) dx_2\end{aligned} \tag{23}$$

Compared with equation (21), two differences are clear. First, the correction of the externality in period 2 includes the reduction in electricity use due to the extra investments. As we have $x_2(\overline{x_1}, I_1(\overline{x_1},\overline{x_2})) < x_2(\overline{x_1})$, this implies that one would overestimate the welfare gain of an actual (i.e., known to be) short-run policy by evaluating equation (21) using estimates of the long-run effect of a policy that was not expected to be short run. Second, the new integral in equation (23) captures the loss in private welfare from these extra investments. One would thus overestimate the welfare effect of a policy that was not expected to be short run by failing to take these investments into account. In fact, the sign of the bias from assuming away hysteresis is ambiguous for such a policy:

$$Bias^{SR}_{NotExp} = \Delta W^{SR}_{NoH} - \Delta W^{SR}_{NotExp} = \beta \int_{x_{20}}^{x_2(\overline{x_1},I_1(\overline{x_1},\overline{x_2}))} MD_2(x_2) dx_2 - \int_{I_1(\overline{x_1})}^{I_1(\overline{x_1},\overline{x_2})} \frac{dV(\overline{x_1},x_2,I_1 I_2)}{dI_1} dI_1 \tag{24}$$

where the welfare effect does not depend on expectations when assuming away hysteresis. The sign of the bias now depends on the sizes of the welfare gain from the correction of the externality in period 2 and of the welfare loss from the extra investments.[67] This latter term is challenging to estimate because one would need to measure the difference in investment choices with and without the wrong policy expectations for all possible types of investments separately, as well as the slope of the demand curve for all these investments. The persistent effect of a policy that was not expected to be short run is thus much less informative.

This discussion highlights the importance of studying *why* agents make persistent changes following short-run policies when evaluating the welfare implications of hysteresis. Note also that this issue only arises when hysteresis is due to costly investments. This highlights the usefulness of distinguishing between the two categories of mechanisms of hysteresis in our framework.

B.5.2 Long-run policy

Next, we show that the persistent effect of a short-run policy also informs the size of the bias from assuming away hysteresis in the case of a long-run policy. The researcher can recover the welfare effect of such a policy by tracing it along any path from (x_{10}, x_{20}) to $(\overline{x_1}, \overline{x_2})$. Yet, with hysteresis,

[67] The welfare effect of this policy could also be larger than that of an actual short-run policy if the gain from the additional correction of the externality outweighs the loss from the extra investments. However, by the *targeting principle* (Sandmo, 1975), a policy in which the government announces and actually implements electricity use restrictions $\overline{x_1}$ and $x_2(\overline{x_1}, I_1(\overline{x_1},\overline{x_2}))$ would lead to the same changes in electricity use for smaller losses in private welfare.

one cannot consider the paths in the two periods separately. Following the path that first changes x_1 to $\overline{x_1}$ and then x_2 to $\overline{x_2}$ given $\overline{x_1}$, which is natural in our intertemporal setting, one obtains:

$$\begin{aligned}
\Delta W^{LR} &= \int_{x_{10}}^{\overline{x_1}} \frac{dW(x_1,x_2,I_1,I_2)}{dx_1} dx_1 + \int_{x_2(\overline{x_1})}^{\overline{x_2}} \frac{dW(\overline{x_1},x_2,I_1 I_2)}{dx_2} dx_2 \\
&= \int_{x_{10}}^{\overline{x_1}} \frac{dV(x_1,x_2,I_1,I_2)}{dx_1} dx_1 + \int_{x_2(\overline{x_1})}^{\overline{x_2}} \frac{dV(\overline{x_1},x_2,I_1,I_2)}{dx_2} dx_2 - \int_{x_{10}}^{\overline{x_1}} \frac{dE_1(x_1)}{dx_1} dx_1 - \beta \int_{x_{20}}^{\overline{x_2}} \frac{dE_2(x_2)}{dx_2} dx_2 \\
&= \int_{x_{10}}^{\overline{x_1}} \frac{dV(x_1,x_2,I_1,I_2)}{dx_1} dx_1 + \beta \int_{x_2(\overline{x_1})}^{\overline{x_2}} \frac{dU_2(\overline{x_1},x_2,I_1(\overline{x_1}),I_2)}{dx_2} dx_2 + \int_{I_1(\overline{x_1})}^{I_1(\overline{x_1},\overline{x_2})} \frac{dV(\overline{x_1},\overline{x_2},I_1,I_2)}{dI_1} dI_1 \\
&\quad - \int_{x_{10}}^{\overline{x_1}} \frac{dE_1(x_1)}{dx_1} dx_1 - \beta \int_{x_{20}}^{\overline{x_2}} \frac{dE_2(x_2)}{dx_2} dx_2 \\
&= \int_{x_{10}}^{\overline{x_1}} [p_1(x_1) - p_{10}]\, dx_1 + \beta \int_{x_2(\overline{x_1})}^{\overline{x_2}} [p_2(x_2|\overline{x_1}, I_1(\overline{x_1})) - p_{20}]\, dx_2 + \int_{I_1(\overline{x_1})}^{I_1(\overline{x_1},\overline{x_2})} \frac{dV(\overline{x_1},\overline{x_2},I_1,I_2)}{dI_1} dI_1 \\
&\quad - \int_{x_{10}}^{\overline{x_1}} MD_1(x_1) dx_1 - \beta \int_{x_{20}}^{\overline{x_2}} MD_2(x_2) dx_2 \qquad (25)
\end{aligned}$$

The welfare gain from correcting the externalities is identical to that in equation (20), so the difference, which we divide into three parts in equation (25), comes from the loss in private welfare. The first integral captures the loss from the period-1 policy, which can still be measured using the demand curve $p_1(x_1)$. This term accounts for changes in private welfare due to changes in x_1 and in the other choice variables to $x_2(\overline{x_1})$, $I_1(\overline{x_1})$, and $I_2(\overline{x_1})$. The second integral captures the loss from the period-2 policy if the household had not anticipated the policy while in period 1. It corresponds to the Harberger triangle under the demand curve $p_2(x_2|\overline{x_1}, I_1(\overline{x_1}))$ and above the baseline price p_{20} between the consumption levels with and without the period-2 policy. This term accounts for changes in private welfare from the additional changes in x_2 to $\overline{x_2}$ and in I_2 to $I_2(\overline{x_1}, \overline{x_2}, I_1(\overline{x_1}))$. Finally, the household could have anticipated the period-2 policy and chosen a different level of investment in period 1. The impact on private welfare is captured by the third integral and corresponds to the change in private welfare from moving the period-1 investment levels from $I_1(\overline{x_1})$ to $I_1(\overline{x_1}, \overline{x_2})$, holding electricity use constant.[68] By revealed preferences, this term is positive because the only reason to make such investments is to mitigate the loss in private welfare due to the period-2 policy.

Equation (8) highlights three sources of bias from assuming away hysteresis:

$$\begin{aligned}
Bias^{LR} &= DWL^{LR}_{NoH} - DWL^{LR} \qquad (26) \\
&= \beta \left[\int_{x_{20}}^{\overline{x_2}} [p_{2,NoH}(x_2) - p_{20}]\, dx_2 - \int_{x_2(\overline{x_1})}^{\overline{x_2}} [p_2(x_2|\overline{x_1}, I_1(\overline{x_1})) - p_{20}]\, dx_2 \right] - \int_{I_1(\overline{x_1})}^{I_1(\overline{x_1},\overline{x_2})} \frac{dV(\overline{x_1},\overline{x_2},I_1,I_2)}{dI_1} dI_1
\end{aligned}$$

First, part of the reduction in electricity use in period 2, from x_{20} to $x_2(\overline{x_1})$, is due to the period-1

[68] The period-2 investment levels endogenously change from $I_2(\overline{x_1}, \overline{x_2}, I_1(\overline{x_1}))$ to $I_2(\overline{x_1}, \overline{x_2})$.

policy. The associated change in private welfare is already accounted for in the period-1 demand curve and should not be double counted.[69] This is why the second integral in equation (25) is only taken from $x_2(\overline{x_1})$ to $\overline{x_2}$. In Figure B.1b, it corresponds to the triangle $A_2D_2E_2$ under the demand curve $p_2(x_2|\overline{x_1}, I_1(\overline{x_1}))$ with base $x_2(\overline{x_1}) - \overline{x_2}$. The first source of bias from assuming away hysteresis corresponds to the trapezoid $D_2B_2C_2E_2$, the difference with the triangle $A_2B_2C_2$ under the demand curve $p_{2,NoH}(x_2)$ with base $x_{20} - \overline{x_2}$. The persistent effect of the short-run policy, $x_{20} - x_2(\overline{x_1})$, which again corresponds to the persistent effect of a policy that was expected to be short run, is thus a key empirical object for evaluating this source of bias.

Second, the change in the period-2 propensity to consume due to the period-1 policy may have changed the slope of the period-2 demand curve. Whether this leads to a bias depends on the variation used to estimate the slopes of the demand curves $p_{2,NoH}(x_2)$ and $p_2(x_2|\overline{x_1}, I_1(\overline{x_1}))$. A researcher who allows for hysteresis would estimate the slope of $p_2(x_2|\overline{x_1}, I_1(\overline{x_1}))$ using exogenous price variation in period 2 following a short-run policy. A researcher who assumes away hysteresis could estimate the slope of $p_{2,NoH}(x_2)$ using the same variation. There would be no second source of bias in this case. However, one may also wrongly believe that the slopes of the demand curves in periods 1 and 2 are identical and use an estimate of the slope of $p_{1,NoH}(x_1)$ for $p_{2,NoH}(x_2)$ as well. The welfare loss would be smaller (resp. larger) in that case, and the bias larger (resp. smaller), if the demand curve became more elastic (resp. inelastic) following the short-run policy, e.g., if households started paying more attention to their electricity use, like in Jessoe and Rapson (2014).

Third, a researcher who assumes away hysteresis would fail to recognize that the household could make different investment choices in period 1 in order to mitigate its overall loss in private welfare.[70] This highlights once again the usefulness of distinguishing between the two mechanisms of hysteresis in our framework. The third source of bias is challenging to estimate. First, one would need to measure the difference in period-1 investments, for all possible types of investments, in the case of a long-run policy and in the case of a short-run policy. Second, one would need to evaluate the change in private welfare resulting from such changes in investments, given electricity use levels $\overline{x_1}$ and $\overline{x_2}$.[71] However, given the sign of this third source of bias, we can abstract from it and still draw meaningful welfare conclusions. In so doing, we evaluate informative bounds: a lower bound for the overall bias and the welfare effect, and an upper bound for the loss in private welfare.

[69] This result carries over to behaviors that are complement within a time period. Suppose, for instance, that smoking and drinking generate negative externalities and that the household likes to smoke when drinking. A policy that reduces drinking would thus also reduce smoking. The demand curve for drinking captures the associated utility from smoking and so the slope of the demand curve for drinking is sufficient to measure any change in utility from the associated change in smoking. Now suppose that another policy also reduces smoking further. The reduction in smoking that took place because of the reduction in drinking should not be double-counted.

[70] This third source of bias is absent for two behaviors that are complements within a given period; there is no room for anticipations to matter when the two behaviors are taking place at the same time.

[71] This last step is conceptually challenging: it is unclear how to estimate demand curves for investments, holding fixed both present and future electricity use levels.

We focus on the first source of bias in our application, which is a natural first step. When we evaluate the size of the bias, we thus assume that a researcher assuming away hysteresis would use the same estimate of the slope of the period-2 demand curve as a researcher allowing for hysteresis. Yet, we acknowledge that our estimates of the welfare effect and the bias may constitute a lower bound if hysteresis is partly due to active investments in the household's propensity to consume.

Note that the key empirical object to evaluate the first source of bias is again the persistent effect of a short-run policy that was expected to be short run. The second integral in equation (25) would be taken over an interval that is too small if one was instead using an estimate of the persistent effect of a policy that was not expected to be short run, as we have: $x_2\left(\overline{x_1}, I_1\left(\overline{x_1}, \overline{x_2}\right)\right) \leq x_2\left(\overline{x_1}\right)$. We can in fact re-write the formulas for the welfare effect and the bias in terms of such an estimate:

$$
\begin{aligned}
\Delta W^{LR} = & \int_{x_{10}}^{\overline{x_1}} \frac{dV(x_1, x_2, I_1, I_2)}{dx_1} dx_1 + \int_{I_1(\overline{x_1})}^{I_1(\overline{x_1}, \overline{x_2})} \frac{dV(\overline{x_1}, x_2, I_1 I_2)}{dI_1} dI_1 \\
& + \beta \int_{x_2(\overline{x_1}, I_1(\overline{x_1}, \overline{x_2}))}^{\overline{x_2}} \frac{dU_2(\overline{x_1}, x_2, I_1(\overline{x_1}, \overline{x_2}), I_2)}{dx_2} dx_2 - \int_{x_{10}}^{\overline{x_1}} \frac{dE_1(x_1)}{dx_1} dx_1 - \beta \int_{x_{20}}^{\overline{x_2}} \frac{dE_2(x_2)}{dx_2} dx_2 \\
= & \int_{x_{10}}^{\overline{x_1}} \left[p_1(x_1) - p_{10}\right] dx_1 + \int_{I_1(\overline{x_1})}^{I_1(\overline{x_1}, \overline{x_2})} \frac{dV(\overline{x_1}, x_2, I_1 I_2)}{dI_1} dI_1 \\
& + \beta \int_{x_2(\overline{x_1}, I_1(\overline{x_1}, \overline{x_2}))}^{\overline{x_2}} \left[p_2(x_2 | \overline{x_1}, I_1(\overline{x_1}, \overline{x_2})) - p_{20}\right] dx_2 - \int_{x_{10}}^{\overline{x_1}} MD_1(x_1) dx_1 - \beta \int_{x_{20}}^{\overline{x_2}} MD_2(x_2) dx_2 \qquad (27)
\end{aligned}
$$

$$
\begin{aligned}
Bias^{LR} = & DWL_{NoH}^{LR} - DWL^{LR} = \beta \left[\int_{x_{20}}^{\overline{x_2}} \left[p_{2,NoH}(x_2) - p_{20}\right] dx_2 - \int_{x_2(\overline{x_1}, I_1(\overline{x_1}, \overline{x_2}))}^{\overline{x_2}} \left[p_2(x_2 | \overline{x_1}, I_1(\overline{x_1}, \overline{x_2})) - p_{20}\right] dx_2 \right] \\
& - \int_{I_1(\overline{x_1})}^{I_1(\overline{x_1}, \overline{x_2})} \frac{dV(\overline{x_1}, x_2, I_1 I_2)}{dI_1} dI_1 \qquad (28)
\end{aligned}
$$

The total loss in private welfare can still be divided into three parts. The first part capturing the loss in private welfare from the period-1 policy is the same as in equation (25). The second part is the loss from the additional investments made in period 1 in preparation for the period-2 policy. The third part is the additional loss from the period-2 policy incurred in period 2. The latter can be estimated using estimates of the long-run effect of a policy that was not expected to be short run and the slope of the demand curve in period 2 after the policy was cancelled at the start of period 2. The first and the third terms can potentially be estimated empirically, but the term involving the change in investment levels remains hard to estimate. The issue is that, abstracting from this term, one would now *overestimate* instead of *underestimate* the welfare effect and the size of the bias. This is less informative because we have limited evidence to begin with that such a bias exists.

Finally, because the correction of the externalities is identical in equations (20) and (25), the bias from assuming away hysteresis is the same whether the government has a short-run or a long-run goal. It is also straightforward from equations (21) and (25) that, even with hysteresis, a government with a short-run (resp. long-run) goal will want to implement a short-run (resp.

long-run) policy.

B.6 Extensions

We now present extensions to the benchmark framework and show that the long-run effect of a short-run policy remains informative for the welfare evaluation of corrective policies.

B.6.1 Adding heterogeneity

First, we can add heterogeneity in our framework and the main results carry through.

The main analysis relies on a representative-household model, but it can be extended to the case of heterogeneous households. One can apply the same arguments as above separately for each household if policies aim at correcting households' behaviors individually. Therefore, in this section, we consider instead policies that aim at *aggregate* changes in behaviors, for instance, when the relevant externalities are functions of aggregate levels of a behavior.

Formally, assume J households facing a problem similar to that of the representative household in the benchmark framework. The solution for each household satisfies the same first-order conditions (17) and (18), with all functions and variables indexed by $j = 1, ..., J$. Define X_1 and X_2 as the aggregate levels of electricity use in the two periods. Without government intervention, the first-order conditions and the baseline electricity prices, p_{10} and p_{20}, will determine the baseline levels of electricity use x^j_{10} and x^j_{20} for each household j and the aggregate levels X_{10} and X_{20}.

Short-run policy. Consider first the short-run policy, which is now defined in terms of aggregate levels, $\overline{X_1} < X_{10}$. Define the vector $\widetilde{x}_1 = (\widetilde{x}^1_1, ..., \widetilde{x}^J_1)$ such that:

$$\sum_j \widetilde{x}^j_1 = \overline{X_1}$$

$$\frac{\partial v^j_1\left(\widetilde{x}^j_1, s^j_1\right)}{\partial \widetilde{x}^j_1} + \beta \frac{\partial v^j_2\left(x^j_2, s^j_2\right)}{\partial s^j_2}\frac{\partial s^j_2}{\partial \widetilde{x}^j_1} = \widetilde{k}_1 \quad ; \quad \frac{\partial v_2\left(x^j_2, s^j_2\right)}{\partial x^j_2} = p_2$$

$$\left[\frac{\partial v^j_1\left(\widetilde{x}^j_1, s^j_1\right)}{\partial s^j_1} + \beta \frac{\partial v^j_2\left(x^j_2, s^j_2\right)}{\partial s^j_2}\frac{\partial s^j_2}{\partial s^j_1}\right]\frac{\partial s^j_1}{\partial I^j_1} - \kappa^{j\prime}_1\left(I^j_1\right) = 0 \quad ; \quad \frac{\partial v^j_2\left(x^j_2, s^j_2\right)}{\partial s^j_2}\frac{\partial s^j_2}{\partial I^j_2} - \kappa^{j\prime}_2\left(I^j_2\right) = 0$$

for each household j and some constant $\widetilde{k}_1$. Define also the vectors $x_2(\widetilde{x}_1)$, $I_1(\widetilde{x}_1)$, and $I_2(\widetilde{x}_1)$ as the optimal choices of consumption levels in period 2 and investment levels in both periods that satisfy the above conditions. The vector $\widetilde{x}_1$ constitutes an efficient allocation of the aggregate quantity $\overline{X_1}$ as marginal returns on investments and marginal utilities of consumption are equal across all agents. This vector would correspond to the market allocation $x_{10} = (x^1_{10}, ..., x^J_{10})$ with $\overline{X_1} = X_{10}$ (and thus $\widetilde{k}_1 = p_{10}$). We consider in this section the welfare effect of short-run policies that are implemented efficiently, for example, with tradable quotas or taxes. Such policies equalize the marginal costs of reducing electricity use across households, therefore minimizing the loss in private welfare with heterogenous households. Moreover, changes in private welfare can be traced along aggregate demand curves for electricity when electricity use is allocated efficiently. Otherwise, allocative inefficiencies arise, complicating the welfare evaluation whether or not one allows for hysteresis.

A researcher who assumes away hysteresis can recover the change in welfare by tracing it along the path from x_{10} to $\widetilde{x}_1$. As before, the researcher would recognize that investment levels in period 1 may adjust endogenously, but she would assume that variables in period 2 remain unaffected:

$$
\begin{aligned}
\Delta W^{SR}_{NoH} &= \sum_j \int_{x^j_{10}}^{\widetilde{x}^j_1} \frac{dV^j(x^j_1, x^j_2, I^j_1, I^j_2)}{dx^j_1} dx^j_1 - \int_{X_{10}}^{\overline{X_1}} \frac{dE_1(X_1)}{dX_1} dX_1 \\
&= \sum_j \int_{x^j_{10}}^{\widetilde{x}^j_1} \frac{dU^j_1(x^j_1, I^j_1)}{dx^j_1} dx^j_1 - \int_{X_{10}}^{\overline{X_1}} \frac{dE_1(X_1)}{dX_1} dX_1 \\
&= \int_{X_{10}}^{\overline{X_1}} \left[P_{1,NoH}(X_1) - p_{10}\right] dX_1 - \int_{X_{10}}^{\overline{X_1}} MD_1(X_1) dX_1 \qquad (29)
\end{aligned}
$$

where we assume that the externality is now a function of the aggregate level of electricity use.

Equation (29) is the aggregate version of the textbook formula for the welfare effect of a corrective policy. The private welfare loss is captured by the area under the aggregate demand curve $P_{1,NoH}(X_1)$ as long as the corrective policy is implemented efficiently. This is because efficient allocations are those that can be traced along aggregate demand curves.

A researcher who allows for the possibility of hysteresis can recover the change in welfare similarly, but she will recognize that x_2 and I_2 may also adjust endogenously:

$$
\begin{aligned}
\Delta W^{SR} &= \sum_j \int_{x^j_{10}}^{\widetilde{x}^j_1} \frac{dV^j(x^j_1, x^j_2, I^j_1, I^j_2)}{dx^j_1} dx^j_1 - \int_{X_{10}}^{\overline{X_1}} \frac{dE_1(X_1)}{dX_1} dX_1 - \beta \int_{X_{20}}^{X_2(\widetilde{x}_1)} \frac{dE_2(X_2)}{dX_2} dX_2 \\
&= \int_{X_{10}}^{\overline{X_1}} \left[P_1(X_1) - p_{10}\right] dX_1 - \int_{X_{10}}^{\overline{X_1}} MD_1(X_1) dX_1 - \beta \int_{X_{20}}^{X_2(\widetilde{x}_1)} MD_2(X_2) dX_2 \qquad (30)
\end{aligned}
$$

where $X_2(\widetilde{x}_1)$ is the aggregate level of electricity use in period 2 given the allocation $\widetilde{x}_1$ in period 1, or the aggregate long-run effect of the short-run policy. It is a function of the allocation, $\widetilde{x}_1$, rather than of the aggregate level, $\overline{X_1}$, because different allocations of the same aggregate level $\overline{X_1}$ in period 1 may result in different aggregate levels X_2 in period 2 with heterogeneous households.

Equation (30) is the aggregate version of the formula in equation (6). The source of the bias from assuming away hysteresis comes again from neglecting the possible correction of the externality in period 2. The new key empirical object to estimate to evaluate this bias is the aggregate long-run effect of the short-run policy, $X_{20} - X_2(\widetilde{x}_1)$, as a researcher would again estimate the demand curve $P_1(X_1)$ in the same way as she would estimate $P_{1,NoH}(X_1)$ if assuming away hysteresis.

Long-run policy. Let's now consider the long-run corrective policy ($\overline{X_1} < X_{10}$, $\overline{X_2} < X_{20}$). Define the vectors $\widehat{x}_1 = (\widehat{x}_1^1, ..., \widehat{x}_1^J)$ and $\widehat{x}_2 = (\widehat{x}_2^1, ..., \widehat{x}_2^J)$ such that:

$$\sum_j \widehat{x}_1^j = \overline{X_1} \quad ; \quad \sum_j \widehat{x}_2^j = \overline{X_2}$$

$$\frac{\partial v_1^j\left(\widehat{x}_1^j, s_1^j\right)}{\partial \widehat{x}_1^j} + \beta \frac{\partial v_2^j\left(\widehat{x}_2^j, s_2^j\right)}{\partial s_2^j}\frac{\partial s_2^j}{\partial \widehat{x}_1^j} = \widehat{k}_1 \quad ; \quad \frac{\partial v_2\left(\widehat{x}_2^j, s_2^j\right)}{\partial \widehat{x}_2^j} = \widehat{k}_2$$

$$\left[\frac{\partial v_1^j\left(\widehat{x}_1^j, s_1^j\right)}{\partial s_1^j} + \beta \frac{\partial v_2^j\left(\widehat{x}_2^j, s_2^j\right)}{\partial s_2^j}\frac{\partial s_2^j}{\partial s_1^j}\right]\frac{\partial s_1^j}{\partial I_1^j} - \kappa_1^{j\prime}\left(I_1^j\right) = 0 \quad ; \quad \frac{\partial v_2^j\left(\widehat{x}_2^j, s_2^j\right)}{\partial s_2^j}\frac{\partial s_2^j}{\partial I_2^j} - \kappa_2^{j\prime}\left(I_2^j\right) = 0$$

for each household j and some constants $\widehat{k}_1$ and $\widehat{k}_2$. Define also the vectors $I_1(\widehat{x}_1, \widehat{x}_2)$, and $I_2(\widehat{x}_1, \widehat{x}_2)$ as the optimal choices of investment levels in both periods that satisfy the above conditions. The vectors $\widehat{x}_1$ and $\widehat{x}_2$ constitute an efficient allocation of the aggregate quantities $\overline{X_1}$ and $\overline{X_2}$. These vectors would correspond to the market allocations $x_{10} = (x_{10}^1, ..., x_{10}^J)$ and $x_{20} = (x_{20}^1, ..., x_{20}^J)$ with $\overline{X_1} = X_{10}$ and $\overline{X_2} = X_{20}$ (and thus $\widehat{k}_1 = p_{10}$ and $\widehat{k}_2 = p_{20}$). For the same reasons as above, we consider again only the welfare effect of long-run corrective policies that are implemented efficiently.

A researcher who assumes away hysteresis can still recover the change in welfare by tracing it along the paths from x_{10} to $\widehat{x}_1$ and from x_{20} to $\widehat{x}_2$ separately, letting I_1 and I_2 adjust endogenously, because behaviors in periods 1 and 2 are assumed to be independent:

$$\begin{aligned}\Delta W_{NoH}^{LR} &= \sum_j \int_{x_{10}^j}^{\widehat{x}_1^j} \frac{dV^j(x_1^j, x_2^j, I_1^j, I_2^j)}{dx_1^j} dx_1^j + \sum_j \int_{x_{20}^j}^{\widehat{x}_2^j} \frac{dV^j(x_1^j, x_2^j, I_1^j, I_2^j)}{dx_2^j} dx_2^j - \int_{X_{10}}^{\overline{X_1}} \frac{dE_1(X_1)}{dX_1} dX_1 - \beta \int_{X_{20}}^{\overline{X_2}} \frac{dE_2(X_2)}{dX_2} dX_2 \\ &= \int_{X_{10}}^{\overline{X_1}} \left[P_{1,NoH}(X_1) - p_{10}\right] dX_1 + \beta \int_{X_{20}}^{\overline{X_2}} \left[P_{2,NoH}(X_2) - p_{20}\right] dX_2 - \int_{X_{10}}^{\overline{X_1}} MD_1(X_1) dX_1 - \beta \int_{X_{20}}^{\overline{X_2}} MD_2(X_2) dX_2\end{aligned} \tag{31}$$

Equation (31) is simply the aggregate version of the two-period textbook formula. Again, the private welfare loss would be captured by the area under the aggregate demand curves in the two periods as long as the policy is assumed to be implemented efficiently.

A researcher who allows for the possibility of hysteresis can still recover the change in welfare by tracing it along any path from (x_{10}, x_{20}) to $(\widehat{x}_1, \widehat{x}_2)$. However, as before, she can no longer consider the paths from x_{10} to $\widehat{x}_1$ and from x_{20} to $\widehat{x}_2$ separately. Moreover, she must take into account the fact that the efficient allocation of $\overline{X_1}$ when there is only a short-run policy, $\widetilde{x}_1$, may be different from the efficient allocation of the same quantity $\overline{X_1}$ in the case of a long-run corrective policy, $\widehat{x}_1$. To derive a formula for the welfare effect, it is then again useful to consider three steps. The first step captures the loss in private welfare from the policy in period 1, as if there was only a short-run policy. The associated change in welfare can be traced along the path from x_{10} to $\widetilde{x}_1$ as for the short-run policy, and it can be measured by the area under the aggregate demand curve $P_1(X_1)$. The second step captures the loss in private welfare from the policy in period 2, if we allocate the quantity $\overline{X_2}$ efficiently, but holding fixed the allocation $\widetilde{x}_1$ and investment levels $I_1(\widetilde{x}_1)$ from period 1. Define the vector $\check{x}_2$, such that:

$$x_1^j = \widetilde{x}_1^j \quad ; \quad I_1^j = I_1(\widetilde{x}_1^j) \quad ; \quad \sum_j \check{x}_2^j = \overline{X_2}$$
$$\frac{\partial v_2\left(\check{x}_2^j, s_2^j\right)}{\partial \check{x}_2^j} = \check{k}_2 \quad ; \quad \frac{\partial v_2^j\left(\check{x}_2^j, s_2^j\right)}{\partial s_2^j}\frac{\partial s_2^j}{\partial I_2^j} = \kappa_2^{j\prime}\left(I_2^j\right)$$

for each household j and some constant $\check{k}_2$. Define also the vector $I_2(\check{x}_2)$ as the optimal choices of investment levels in period 2 that satisfy the above conditions. The change in private welfare in this second step can then be traced along the path from $x_2(\widetilde{x}_1)$ to $\check{x}_2$ under an aggregate demand curve $P_2(X_2|\widetilde{x}_1, I_1(\widetilde{x}_1))$ estimated using price variation once in period 2, following a short-run policy. Once again, a researcher assuming away hysteresis would use the same variation to estimate $P_{2,NoH}(X_2)$. Finally, the last step captures the fact that the allocations $\widetilde{x}_1$ and $\check{x}_2$, and the investment levels $I_1(\widetilde{x}_1)$ and $I_2(\check{x}_2)$ may not correspond to the allocations and investment levels that minimize the overall loss in private welfare – the vectors $\widehat{x}_1$, $\widehat{x}_2$, $I_1(\widehat{x}_1, \widehat{x}_2)$, and $I_2(\widehat{x}_1, \widehat{x}_2)$. This last step thus (weakly) reduces the loss in private welfare from the first two steps. Formally, we have:

$$\begin{aligned}
\Delta W^{LR} &= \sum_j \int_{x_{10}^j}^{\widetilde{x}_1^j} \frac{dV^j(x_1^j, x_2^j, I_1^j, I_2^j)}{dx_1^j} dx_1^j + \beta \sum_j \int_{x_2^j(\widetilde{x}_1^j)}^{\check{x}_2^j} \frac{dU_2^j(\widetilde{x}_1^j, x_2^j, I_1^j(\widetilde{x}_1^j), I_2^j)}{dx_2^j} dx_2^j \\
&+ \sum_j \left[V^j(\widehat{x}_1^j, \widehat{x}_2^j, I_1^j(\widehat{x}_1^j, \widehat{x}_2^j), I_2^j(\widehat{x}_1^j, \widehat{x}_2^j)) - V^j(\widetilde{x}_1^j, \check{x}_2^j, I_1^j(\widetilde{x}_1^j), I_2^j(\check{x}_2^j))\right] \\
&- \int_{X_{10}}^{\overline{X_1}} \frac{dE_1(X_1)}{dX_1} dX_1 - \beta \int_{X_{20}}^{\overline{X_2}} \frac{dE_2(X_2)}{dX_2} dX_2 \\
&= \int_{X_{10}}^{\overline{X_1}} \left[P_1(X_1) - p_{10}\right] dX_1 + \beta \int_{X_2(\widetilde{x}_1)}^{\overline{X_2}} \left[P_2(X_2|\widetilde{x}_1, I_1(\widetilde{x}_1)) - p_{20}\right] dX_2 \\
&+ \sum_j \left[V^j(\widehat{x}_1^j, \widehat{x}_2^j, I_1^j(\widehat{x}_1^j, \widehat{x}_2^j), I_2^j(\widehat{x}_1^j, \widehat{x}_2^j)) - V^j(\widetilde{x}_1^j, \check{x}_2^j, I_1^j(\widetilde{x}_1^j), I_2^j(\check{x}_2^j))\right] \\
&- \int_{X_{10}}^{\overline{X_1}} MD_1(X_1) dX_1 - \beta \int_{X_{20}}^{\overline{X_2}} MD_2(X_2) dX_2
\end{aligned} \tag{32}$$

Equation (32) highlights the same three potential sources of bias when assuming away hysteresis as in equation (25). First, the researcher would fail to recognize that part of the reduction in electricity use in period 2 may be the result of the period-1 policy. The key empirical object to estimate to evaluate this bias is the aggregate long-run effect of the short-run policy, $X_{20} - X_2(\widetilde{x}_1)$. Second, there may be a bias if the researcher uses the wrong slope for the demand curve in period 2. Third, there are gains from anticipating the period-2 policy. The third source of bias is even more difficult to estimate with heterogeneous agents. It still captures the fact that mechanisms of hysteresis allow households to make different investment choices in order to mitigate the overall private welfare loss from the long-run corrective policy. However, with heterogeneous agents, it also captures the fact that mechanisms of hysteresis provide new gains from trade across households (moving from the allocation $\check{x}_2$ to the allocation $\widehat{x}_2$), which can further mitigate the overall loss in private welfare from the long-run corrective policy. One would thus further underestimate the overall bias by abstracting from the third source of bias in the case of heterogeneous households.

B.6.2 Adding uncertainty

The benchmark framework abstracts from any uncertainty in the household decision-making. However, a lot of choices involve uncertain costs and benefits, for instance when deciding to adopt a new technology or behavior. Moreover, resolving such uncertainty before making decisions may be costly. This is why scholars have suggested that models of rational inattention may help describe household decisionmaking, including for energy-related choices (e.g., Sallee, 2014). Furthermore, to the extent that the knowledge acquired by resolving uncertainty does not depreciate immediately, it constitutes a possible mechanism of hysteresis. We thus add uncertainty by considering a rational inattention version of the benchmark framework, and we show that our main takeaways survive.[72]

We assume that the household faces the same problem as in the benchmark framework, but that it is now uncertain about the cost of investing in its propensity to consume:

$$
\begin{aligned}
V(\theta) =& y_1 - p_1 x_1 - \theta \kappa_1 (I_1) + \nu_1 (x_1, s_1) + \beta \left[y_2 - p_2 x_2 - \theta \kappa_2 (I_2) + \nu_2 (x_2, s_2) \right] \\
\text{s.t. } s_t =& s_t (s_{t-1}, x_{t-1}, I_t), \text{ for } t = 1,2 \text{ and } s_0, x_0 \text{ given}
\end{aligned} \tag{33}
$$

where $\theta \in [\underline{\theta}, \overline{\theta}]$ is a random variable with mean $E[\theta]$. We assume that the household knows the distribution of θ, but doesn't know the actual realization of the random variable before making choices, unless it devotes some costly "attention" effort. In particular, prior to the start of period 1, the household can decide how much effort e, with strictly convex cost $\psi(e)$ to devote to learn about the realization of θ. With probability e, the uncertainty is then resolved for both periods (the household learns the value of θ); otherwise, the household make choices under uncertainty. This assumption allows the resolution of the uncertainty to be a mechanism of hysteresis: the household may adjust its attention effort in response to a corrective policy in period 1, and changes in the information it acquires in the process may affect its choices persistently.

[72] In such a model, alternative policies such as information campaigns (i.e., decreasing the cost of information acquisition) may become useful policy instruments as well.

Formally, the household solves:

$$Z = max_e \quad e \cdot E[max_{x_1,I_1,x_2,I_2} V(\theta)] + (1-e) \cdot max_{x_1,I_1,x_2,I_2} E[V(\theta)] - \psi(e)$$

The first-order conditions for this problem are:

$$E[max_{x_1,I_1,x_2,I_2}(U_1(\theta)+\beta U_2)] - max_{x_1,I_1,x_2,I_2} E[U_1(\theta)+\beta U_2] - \psi'(e) = 0 \tag{34}$$

$$\text{if learn } \theta: \quad \frac{\partial v_1(x_1,s_1)}{\partial x_1} + \beta \frac{\partial v_2(x_2,s_2)}{\partial s_2}\frac{\partial s_2}{\partial x_1} = p_1 \quad ; \quad \frac{\partial v_2(x_2,s_2)}{\partial x_2} = p_2$$
$$\left[\frac{\partial v_1(x_1,s_1)}{\partial s_1} + \beta \frac{\partial v_2(x_2,s_2)}{\partial s_2}\frac{\partial s_2}{\partial s_1}\right]\frac{\partial s_1}{\partial I_1} - \theta \kappa_1'(I_1) = 0 \quad ; \quad \frac{\partial v_2(x_2,s_2)}{\partial s_2}\frac{\partial s_2}{\partial I_2} - \theta \kappa_2'(I_2) = 0 \tag{35}$$

$$\text{otherwise:} \quad \frac{\partial v_1(x_1,s_1)}{\partial x_1} + \beta \frac{\partial v_2(x_2,s_2)}{\partial s_2}\frac{\partial s_2}{\partial x_1} = p_1 \quad ; \quad \frac{\partial v_2(x_2,s_2)}{\partial x_2} = p_2$$
$$\left[\frac{\partial v_1(x_1,s_1)}{\partial s_1} + \beta \frac{\partial v_2(x_2,s_2)}{\partial s_2}\frac{\partial s_2}{\partial s_1}\right]\frac{\partial s_1}{\partial I_1} - E[\theta]\kappa_1'(I_1) = 0 \quad ; \quad \frac{\partial v_2(x_2,s_2)}{\partial s_2}\frac{\partial s_2}{\partial I_2} - E[\theta]\kappa_2'(I_2) = 0 \tag{36}$$

The household will still reduce its electricity use and invest more in its propensity to consume in period 1 if its period-1 choices affect the propensity to consume in period 2, whether or not it learns about the realization θ. There will be three effects of an increase in prices in this model: (i) the household will directly reduce its consumption of electricity; (ii) the household will invest in its propensity to consume, indirectly reducing its consumption of electricity; (iii) the household may change the attention effort it devotes to resolving the uncertainty around θ, also indirectly affecting its investment levels and thus its consumption of electricity. In models of rational inattention, the latter effect may be non-monotonic: at low electricity prices, the household may not need to invest much in its propensity to consume even if it is relatively cheap, so the value of resolving the uncertainty around θ may be low; at higher prices, the household may be interested in making investments depending on their costs, raising the value of information and the effort that the household is willing to devote to resolving the uncertainty around θ; at very high prices, the household may decide to make large investments even without knowing the value of θ, so the value of resolving the uncertainty around θ may again be low. These three effects will still give rise to well-behaved demand curves for electricity as long as they evolve smoothly with changes in prices.

Welfare must now take into account attention costs, so we have: $W = Z - E_1(x_1) - \beta E_2(x_2)$. The welfare gain from corrective policies will still arise from the correction of the externality and the welfare loss from the change in private welfare from pushing the household to reduce electricity use below privately optimal levels, but the latter will now include changes in private welfare due to changes in attention efforts.

Short-run policy. Consider first the short-run policy ($\overline{x_1} < x_{10}$). A researcher who assumes away hysteresis must assume that any resolution of the uncertainty in period 1 has no consequences in period 2. Therefore, the household must be assumed to resolve the uncertainty separately for periods 1 and 2, such as in the following:

$$
\begin{aligned}
Z_{NoH} = max_{e_1,e_2} \quad & e_1 \cdot E[max_{x_1,I_1} U_1(\theta_1)] + (1-e_1) \cdot max_{x_1,I_1} E[U_1(\theta_1)] - \psi(e_1) \\
& + \beta \left[e_2 \cdot E[max_{x_2,I_2} U_2(\theta_2)] + (1-e_2) \cdot max_{x_2,I_2} E[U_2(\theta_2)] - \psi(e_2) \right]
\end{aligned}
$$

In that case, the researcher can recover the change in welfare by tracing it along the path from x_{10} to $\overline{x_1}$, letting I_1 but also e_1 adjust endogenously, and assuming that variables in period 2 are unaffected:

$$
\begin{aligned}
\Delta W^{SR}_{NoH} &= \int_{x_{10}}^{\overline{x_1}} \frac{dZ_{NoH}(e_1, e_2, x_1^\theta, x_2^\theta, I_1^\theta, I_2^\theta)}{dx_1} dx_1 - \int_{x_{10}}^{\overline{x_1}} \frac{dE_1(x_1)}{dx_1} dx_1 \\
&= \int_{x_{10}}^{\overline{x_1}} \left[p_{1,NoH}(x_1) - p_{10} \right] dx_1 - \int_{x_{10}}^{\overline{x_1}} MD_1(x_1) dx_1 \qquad (37)
\end{aligned}
$$

Equation (37) is again the textbook formula. The slope of the demand curve in period 1 remains a sufficient statistic, because it captures the value that the household attach to a unit of electricity use, given its possibility to adapt endogenously its investment and information acquisition choices.

A researcher who allows for hysteresis can also recover the change in welfare by tracing it along the path from x_{10} to $\overline{x_1}$, letting the other variables, including x_2, adjust endogenously:

$$
\begin{aligned}
\Delta W^{SR} &= \int_{x_{10}}^{\overline{x_1}} \frac{dZ(e, x_1^\theta, x_2^\theta, I_1^\theta, I_2^\theta)}{dx_1} dx_1 - \int_{x_{10}}^{\overline{x_1}} \frac{dE_1(x_1)}{dx_1} dx_1 - \beta \int_{x_{10}}^{\overline{x_1}} \frac{dE_2(x_2)}{dx_2} \frac{dx_2}{dx_1} dx_1 \qquad (38) \\
&= \int_{x_{10}}^{\overline{x_1}} \left[p_1(x_1) - p_{10} \right] dx_1 - \int_{x_{10}}^{\overline{x_1}} MD_1(x_1) dx_1 - \beta \int_{x_{20}}^{x_2(\overline{x_1})} MD_2(x_2) dx_2
\end{aligned}
$$

Equation (38) is the same formula as in the benchmark framework because the demand curve in period 1 will take into account changes in private welfare coming from changes in e and x_2.

Long-run policy. Let's now consider the long-run policy ($\overline{x_1} < x_{10}$, $\overline{x_2} < x_{20}$). A researcher who assumes away hysteresis can recover the change in welfare by tracing it from x_{10} to $\overline{x_1}$ and from x_{20} to $\overline{x_2}$ separately, letting I_1, I_2, e_1 and e_2 adjust endogenously:

$$\begin{aligned}
\Delta W_{NoH}^{LR} &= \int_{x_{10}}^{\overline{x_1}} \frac{dZ_{NoH}(e_1,e_2,x_1^\theta,x_2^\theta,I_1^\theta,I_2^\theta)}{dx_1}dx_1 + \beta\int_{x_{20}}^{\overline{x_2}} \frac{dZ_{NoH}(e_1,e_2,x_1^\theta,x_2^\theta,I_1^\theta,I_2^\theta)}{dx_2}dx_2 \\
&\quad - \int_{x_{10}}^{\overline{x_1}} \frac{dE_1(x_1)}{dx_1}dx_1 - \beta\int_{x_{20}}^{\overline{x_2}} \frac{dE_2(x_2)}{dx_2}dx_2 \\
&= \int_{x_{10}}^{\overline{x_1}} [p_{1,NoH}(x_1) - p_{10}]\,dx_1 + \beta\int_{x_{20}}^{\overline{x_2}} [p_{2,NoH}(x_2) - p_{20}]\,dx_2 - \int_{x_{10}}^{\overline{x_1}} MD_1(x_1)dx_1 - \beta\int_{x_{20}}^{\overline{x_2}} MD_2(x_2)dx_2
\end{aligned} \tag{39}$$

Equation (39) is a two-period version of equation (37), with demand curves taking into account endogenous changes in other variables, including attention efforts, along changes in electricity use.

It is useful to consider again three steps to derive a formula for the welfare effect in the case of a researcher who allows for hysteresis. The first step captures the loss in private welfare from the period-1 policy, as if there was only a short-run policy. The second step captures the loss in private welfare from the policy in period 2, holding fixed investment and attention effort levels from period 1. Finally, the third step captures the fact that the household could anticipate the period-2 policy in period 1, and therefore could potentially choose different levels of investment and attention effort. Define $e(\overline{x_1})$, the attention effort chosen in the case of a short-run policy, and $e(\overline{x_1},\overline{x_2})$ the attention effort chosen in the case of a long-run corrective policy. We then have:

$$\begin{aligned}
\Delta W^{LR} &= \int_{x_{10}}^{\overline{x_1}} \frac{dZ(e,x_1^\theta,x_2^\theta,I_1^\theta,I_2^\theta)}{dx_1}dx_1 + \beta\int_{x_2(\overline{x_1})}^{\overline{x_2}} \frac{dZ(e,\overline{x_1},x_2^\theta,I_1^\theta,I_2^\theta)}{dx_2}dx_2 \\
&\quad - \int_{x_{10}}^{\overline{x_1}} \frac{dE_1(x_1)}{dx_1}dx_1 - \beta\int_{x_{20}}^{\overline{x_2}} \frac{dE_2(x_2)}{dx_2}dx_2 \\
&= \int_{x_{10}}^{\overline{x_1}} [p_1(x_1) - p_{10}]\,dx_1 + \beta\int_{x_2(\overline{x_1})}^{\overline{x_2}} [p_2(x_2|\overline{x_1},I_1(\overline{x_1}),e(\overline{x_1})) - p_{20}]\,dx_2 \\
&\quad + [Z(e(\overline{x_1},\overline{x_2}),\overline{x_1},\overline{x_2},I_1(\overline{x_1},\overline{x_2}),I_2(\overline{x_1},\overline{x_2})) - Z(e(\overline{x_1}),\overline{x_1},\overline{x_2},I_1(\overline{x_1}),I_2(e(\overline{x_1}),\overline{x_1},\overline{x_2},I_1(\overline{x_1})))] \\
&\quad - \int_{x_{10}}^{\overline{x_1}} MD_1(x_1)dx_1 - \beta\int_{x_{20}}^{\overline{x_2}} MD_2(x_2)dx_2
\end{aligned} \tag{40}$$

The main difference in equation (40) from the benchmark framework relates again to the third source of bias in the case of the long-run policy. Indeed, the household could have also adjusted its attention effort, e, in period 1 in anticipation of the period-2 policy, further mitigating the loss in its private welfare. One would thus further underestimate the overall bias from assuming away hysteresis by abstracting from the third source of bias, which is difficult to estimate in practice.

B.6.3 Adding a producer surplus

We now relax the assumption that electricity is priced at constant marginal costs, such that welfare includes a producer surplus.[73] Our main results hold in that case for the long-run policy, but the sign of the bias from assuming away hysteresis becomes ambiguous for the short-run policy.

The household's problem is unchanged, but we now assume that electricity is provided by a representative price-taking firm maximizing profits $\pi = p_1 x_1 - c_1(x_1) + \beta\,[p_2 x_2 - c_2(x_2)]$, with strictly increasing and convex costs $c_1(x_1)$ and $c_2(x_2)$. The solution to the firm problem must satisfy the usual first-order conditions: $c_1'(x_1) = p_1$ and $c_2'(x_2) = p_2$. In that model, baseline prices (p_{10}, p_{20}) and quantities (x_{10}, x_{20}) are such that the first-order conditions of both the household and the firm are satisfied. We also consider an alternative model that better fit the regulatory framework in our application, in which marginal costs are constant (i.e., $c_t' = c_t'(x_t) \forall x_t$) but the firm is allowed to charge a fixed mark-up (e.g., to cover some fixed costs). The profit is then $\pi = \mu_1 x_1 + \beta\,[\mu_2 x_2]$, where $\mu_t = p_t - c_t'$ is the fixed markup over constant marginal cost c_t' in period $t = 1,2$. In that model, the baseline prices (p_{10}, p_{20}) are given by $p_{t0} = c_t' + \mu_t$, and the baseline quantities (x_{10}, x_{20}) are determined by the first-order conditions of the household.

Welfare must now take into account the firm profits, so we have: $W = V + \pi - E_1(x_1) - \beta E_2(x_2)$. The welfare gain from corrective policies still only arises from the correction of the externality. However, the welfare loss now comes from changes in both the household utility (i.e., the consumer surplus) and the firm profits (i.e., the producer surplus), when inducing the household to reduce electricity use below baseline levels.

Short-run policy. Consider first the short-run policy ($\overline{x_1} < x_{10}$). Assuming away hysteresis, the researcher can still recover the change in welfare by tracing it along the path from x_{10} to $\overline{x_1}$, letting I_1 adjust endogenously, and assuming that variables in period 2 remain unaffected:

$$\begin{aligned}\Delta W_{NoH}^{SR} &= \int_{x_{10}}^{\overline{x_1}} \frac{dV(x_1,x_2,I_1,I_2)}{dx_1}dx_1 + \int_{x_{10}}^{\overline{x_1}} \frac{d\pi(x_1,x_2)}{dx_1}dx_1 - \int_{x_{10}}^{\overline{x_1}} \frac{dE_1(x_1)}{dx_1}dx_1 \\ &= \int_{x_{10}}^{\overline{x_1}} [p_{1,NoH}(x_1) - p_{10}]\,dx_1 + \int_{x_{10}}^{\overline{x_1}} \left[p_{10} - c_1'(x_1)\right] dx_1 - \int_{x_{10}}^{\overline{x_1}} MD_1(x_1)dx_1 \\ &= \int_{x_{10}}^{\overline{x_1}} \left[p_{1,NoH}(x_1) - c_1'(x_1)\right] dx_1 - \int_{x_{10}}^{\overline{x_1}} MD_1(x_1)dx_1 \end{aligned} \tag{41}$$

Equation (41) is the textbook formula with a producer surplus. The term capturing the correction of the externality is unchanged. However, the Harberger triangle is now the area under the demand curve and above the marginal cost curve (i.e., the supply curve; it is horizontal at $c_1'(x_1) = c_1'$ in the alternative model), between the consumption levels with and without the corrective policy, in the period in which the policy is implemented. This corresponds to the triangle $A_1B_1C_1$ in Figure

[73] The hysteresis still takes place on the demand side, and not on the firm side like in, e.g., Acemoglu et al. (2012).

B.2a for the first model and the trapezoid $A_1G_1B_1C_1$ in Figure B.3a for the alternative model. The only additional empirical object to estimate for evaluating the welfare effect, compared to the benchmark framework, is the slope of the supply curve in period 1 in the first model or the markup in period 1 in the alternative model.

Allowing for hysteresis, the welfare effect of the short-run policy includes the same additional welfare gain as in the benchmark framework from the correction of the externality in period 2, but the hysteresis in electricity use now also implies an additional loss in producer surplus in period 2.

In the case of the alternative model (*AM*), we can still recover the change in welfare by tracing it along the path from x_{10} to $\overline{x_1}$, allowing the other variables to adjust endogenously:

$$\begin{aligned}\Delta W_{AM}^{SR} &= \int_{x_{10}}^{\overline{x_1}} \frac{dV(x_1,x_2,I_1,I_2)}{dx_1}dx_1 + \int_{x_{10}}^{\overline{x_1}} \frac{d\pi(x_1,x_2)}{dx_1}dx_1 - \int_{x_{10}}^{\overline{x_1}} \frac{dE_1(x_1)}{dx_1}dx_1 - \beta\int_{x_{10}}^{\overline{x_1}} \frac{dE_2(x_2)}{dx_2}\frac{dx_2}{dx_1}dx_1 \\ &= \int_{x_{10}}^{\overline{x_1}} \left[p_1(x_1) - c_1'(x_1)\right]dx_1 + \beta\int_{x_{20}}^{x_2(\overline{x_1})} \left[p_{20} - c_2'\right]dx_2 - \int_{x_{10}}^{\overline{x_1}} MD_1(x_1)dx_1 - \beta\int_{x_{20}}^{x_2(\overline{x_1})} MD_2(x_2)dx_2\end{aligned} \quad (42)$$

There are two novelties compared to the benchmark framework. The first integral in equation (42) accounts for the loss in producer surplus in period 1, as when we assume away hysteresis. The second integral accounts for a loss in producer surplus in period 2, from the unrealized profits on the electricity that the firm does not sell in period 2 because of the policy. This corresponds to the rectangle $H_2G_2B_2D_2$ in Figure B.3b. Equation (42) shows that the mark-up in period 2, $p_{20} - c_2'$, must now be estimated in order to evaluate welfare effects.

In this model, the bias from assuming away hysteresis becomes:

$$Bias_{AM}^{SR} = \beta\int_{x_{20}}^{x_2(\overline{x_1})} MD_2(x_2)dx_2 - \beta\int_{x_{20}}^{x_2(\overline{x_1})} \left[p_{20} - c_2'\right]dx_2 \quad (43)$$

The sign of the bias is now ambiguous because of the additional source of welfare loss. In other words, assuming away hysteresis, the researcher will still miscalculate the welfare effect, but she may not actually underestimate it, depending on the relative size of the welfare gain from the correction of the externality in period 2 and of the loss in producer surplus in period 2.

In the case of the first model (*FM*), we must take into account several considerations. First, equilibrium prices will decrease in period 2 if there is hysteresis, and thus the quantity demanded by the household decreases, because of the upward-sloping supply curve. Second, while the demand curve in period 1 factors in changes in the household utility in period 2 resulting from the policy, it will not account for the fact that the equilibrium price will adjust in period 2 for price-taking customers. Third, as with the alternative model, the producer surplus will also change in period 2 if the equilibrium price and quantity decrease. Formally, we can recover the change in welfare in two steps, by first tracing it along the path from x_{10} to $\overline{x_1}$, and then from $x_2(\overline{x_1}, p_{20})$ to $x_2(\overline{x_1}, p_2)$;

this notation refers to consumption levels in period 2 before and after the equilibrium price adjusts.

$$\begin{aligned}\Delta W_{FM}^{SR} =& \int_{x_{10}}^{\overline{x_1}} \frac{dV(x_1,x_2,I_1,I_2)}{dx_1}dx_1 + \int_{x_{10}}^{\overline{x_1}} \frac{d\pi(x_1,x_2)}{dx_1}dx_1 - \int_{x_{10}}^{\overline{x_1}} \frac{dE_1(x_1)}{dx_1}dx_1 - \beta \int_{x_{10}}^{\overline{x_1}} \frac{dE_2(x_2)}{dx_2}\frac{dx_2}{dx_1}dx_1 \\ =& \int_{x_{10}}^{\overline{x_1}} \left[p_1(x_1) - c_1'(x_1)\right] dx_1 + \beta \int_{x_{20}}^{x_2(\overline{x_1},p_{20})} \left[p_{20} - c_2'(x_2)\right] dx_2 + \beta \int_{x_2(\overline{x_1},p_{20})}^{x_2(\overline{x_1},p_2)} \left[p_2(x_2|\overline{x_1},I_1(\overline{x_1})) - c_2'(x_2)\right] dx_2 \\ & - \int_{x_{10}}^{\overline{x_1}} MD_1(x_1)dx_1 - \beta \int_{x_{20}}^{x_2(\overline{x_1},p_2)} MD_2(x_2)dx_2 \end{aligned} \tag{44}$$

There are four novelties compared to the benchmark framework. The first integral in equation (44) accounts for the loss in producer surplus in period 1, as when we assume away hysteresis. The second integral accounts for the loss in producer surplus in period 2 that would take place if the equilibrium price did not adjust. This corresponds to the triangle $B_2D_2G_2$ in Figure B.2b. The third integral accounts for the fact that the equilibrium price will adjust in period 2 increasing both the consumer and producer surpluses.[74] This corresponds to the triangle $D_2F_2G_2$. The sum of these two last terms results in a welfare loss, which corresponds to the triangle $B_2D_2F_2$. Finally, the integral capturing the welfare gain from correcting the externality in period 2 goes from the baseline quantity to the quantity that will actually be observed in period 2, after the equilibrium price adjusts. Equation (44) shows that the slope of the supply curve in period 2 and the quantity that would have been observed in absence of price adjustment, $x_2(\overline{x_1}, p_{20})$, must now be estimated in order to evaluate welfare effects. The latter statistic is not observed in practice. However, it can be recovered using estimates of the slope of the demand curve in period 2, and of the long-term effects of the short-run policy on both quantity and price, $x_2(\overline{x_1}, p_2) - x_{20}$ and $p_2 - p_{20}$.

In this model, the bias from assuming away hysteresis becomes:

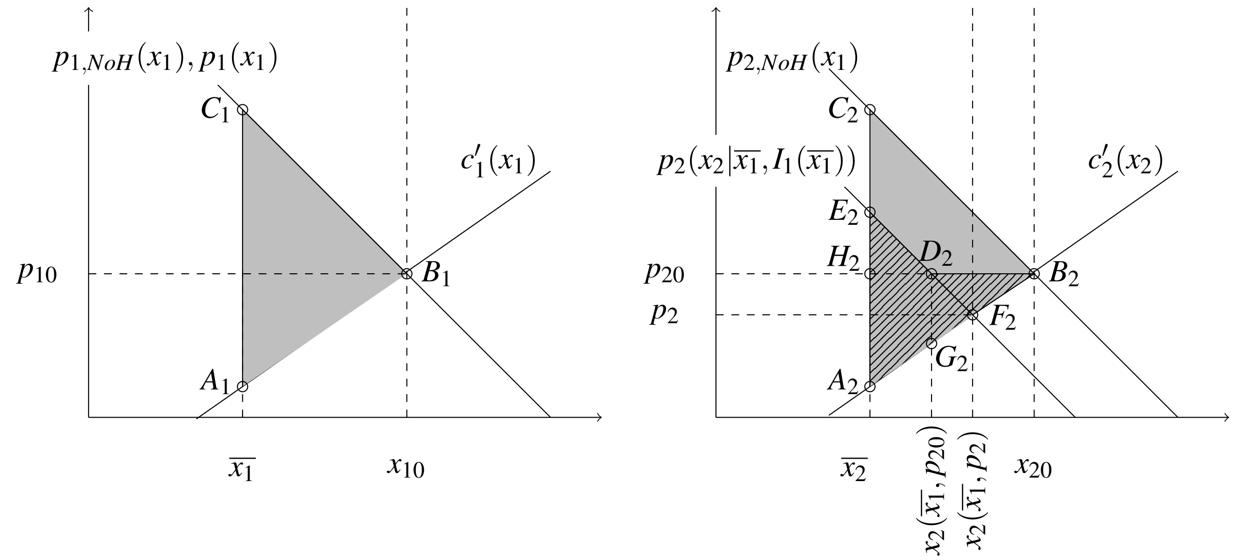
$$Bias^{SR} = \beta \int_{x_{20}}^{x_2(\overline{x_1},p_2)} MD_2(x_2)dx_2 - \left[\beta \int_{x_{20}}^{x_2(\overline{x_1},p_{20})} \left[p_{20} - c_2'(x_2)\right] dx_2 + \beta \int_{x_2(\overline{x_1},p_{20})}^{x_2(\overline{x_1},p_2)} \left[p_2(x_2|\overline{x_1},I_1(\overline{x_1})) - c_2'(x_2)\right] dx_2\right] \tag{45}$$

As in the alternative model, the sign of the bias is ambiguous because of the additional source of welfare loss.

Long-run policy. Let's consider now the long-run policy ($\overline{x_1} < x_{10}, \overline{x_2} < x_{20}$). Assuming away hysteresis, the researcher can still recover the change in welfare by tracing it along the paths from

[74] Technically, we should also write $p_2(x_2|\overline{x_1}, I_1(\overline{x_1}, p_{20}))$ as the household likely makes investment decisions in period 1 without anticipating the change in the equilibrium price in period 2.

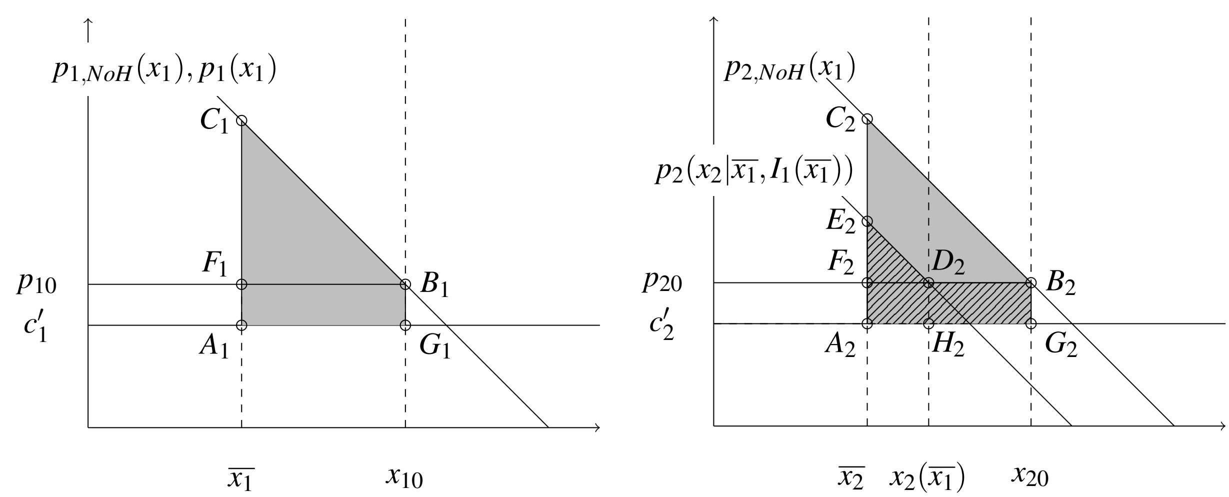
Figure B.2: The loss in private welfare from corrective policies with and without hysteresis in a model with an upward-sloping supply curve ("first model")



(a) Period 1 (b) Period 2

The figure illustrates the loss in private welfare in a model with an upward-sloping supply curve (the private welfare corresponds to the sum of the consumer and producer surpluses) for the short-run and the long-run corrective polices. For simplicity, we assume linear demand curves and no change in the slope of the demand curve following a short-run policy in the presence of hysteresis. See text for details.

Figure B.3: The loss in private welfare from corrective policies with and without hysteresis in a model with a fixed mark-up over constant marginal cost ("alternative model")



(a) Period 1 (b) Period 2

The figure illustrates the loss in private welfare in a model with a fixed mark-up over constant marginal cost (the private welfare corresponds to the sum of the consumer and producer surpluses) for the short-run and the long-run corrective polices. For simplicity, we assume linear demand curves and no change in the slope of the demand curve following a short-run policy in the presence of hysteresis. See text for details.

x_{10} to $\overline{x_1}$ and from x_{20} to $\overline{x_2}$ separately, letting I_1 and I_2 adjust endogenously:

$$\begin{aligned}
\Delta W^{LR}_{NoH} &= \int_{x_{10}}^{\overline{x_1}} \frac{dU_1(x_1,I_1)}{dx_1}dx_1 + \int_{x_{10}}^{\overline{x_1}} \frac{d\pi(x_1,x_2)}{dx_1}dx_1 + \beta \int_{x_{20}}^{\overline{x_2}} \frac{dU_2(x_2,I_2)}{dx_2}dx_2 + \int_{x_{20}}^{\overline{x_2}} \frac{d\pi(x_1,x_2)}{dx_2}dx_2 dx_1 \\
&\quad - \int_{x_{10}}^{\overline{x_1}} \frac{dE_1(x_1)}{dx_1}dx_1 - \beta \int_{x_{20}}^{\overline{x_2}} \frac{dE_2(x_2)}{dx_2}dx_2 \qquad (46) \\
&= \int_{x_{10}}^{\overline{x_1}} \left[p_{1,NoH}(x_1) - c'(x_1)\right] dx_1 + \beta \int_{x_{20}}^{\overline{x_2}} \left[p_{2,NoH}(x_2) - c'(x_2)\right] dx_2 - \int_{x_{10}}^{\overline{x_1}} MD_1(x_1)dx_1 - \beta \int_{x_{20}}^{\overline{x_2}} MD_2(x_2)dx_2
\end{aligned}$$

Equation (46) is simply a two-period version of equation (41). The change in private welfare now corresponds to the sum of the triangles $A_1B_1C_1$ and $A_2B_2C_2$ in Figures B.2a and B.2b for the first model, and of the trapezoids $A_1G_1B_1C_1$ and $A_2G_2B_2C_2$ in Figures B.3a and B.3b for the alternative model. The only additional empirical objects to estimate for evaluating the welfare effect are the slopes of the supply curves in both periods in the first model or the markups in both periods in the alternative model.

It is useful to consider the same three steps as before to derive a formula for the welfare effect when allowing for the possibility of hysteresis. The first step captures the loss in private welfare from the policy in period 1, as if there was only a short-run policy. We do not have to account for any change in equilibrium prices in period 2, as in the case of a short-run policy, because the period-2 quantity will not be determined by the market (assuming $\overline{x_2}$ is binding). The second step captures the loss in private welfare from the period-2 policy, holding fixed period-1 investment levels. Finally, the third step captures the fact that the household could have anticipated the period-2 policy in period 1, and therefore potentially chosen different levels of investment. We then have:

$$\begin{aligned}
\Delta W^{LR} &= \int_{x_{10}}^{\overline{x_1}} \left[p_1(x_1) - c_1'(x_1)\right] dx_1 + \beta \int_{x_{20}}^{x_2(\overline{x_1},p_{20})} \left[p_{20} - c_2'(x_2)\right] dx_2 + \beta \int_{x_2(\overline{x_1},p_{20})}^{\overline{x_2}} \left[p_2(x_2|\overline{x_1}, I_1(\overline{x_1})) - c_2'(x_2)\right] dx_2 \\
&\quad + \int_{I_1(\overline{x_1})}^{I_1(\overline{x_1},\overline{x_2})} \frac{dV(\overline{x_1},\overline{x_2},I_1,I_2)}{dI_1} dI_1 - \int_{x_{10}}^{\overline{x_1}} MD_1(x_1)dx_1 - \beta \int_{x_{20}}^{\overline{x_2}} MD_2(x_2)dx_2 \qquad (47)
\end{aligned}$$

where we have $x_2(\overline{x_1}, p_{20}) = x_2(\overline{x_1})$ and $c_t'(x_t) = c_t'$ for the alternative model. Equation (47) differs from the benchmark framework in three ways. First, it takes into account the loss in producer surplus from the policy in period 1. Second, it takes into account the loss in producer surplus in period 2 due to the hysteresis in electricity demand, $x_2(\overline{x_1}, p_{20}) < x_{20}$. This is again captured by the triangle $B_2D_2G_2$ in Figure B.2b for the first model and by the rectangle $H_2G_2B_2D_2$ in Figure B.3b for the alternative model. Third, it takes into account the loss in producer surplus for the additional reduction in quantity due to the period-2 policy. This is captured by the trapezoid $A_2G_2D_2H_2$ in Figure B.2b for the first model and by the rectangle $A_2H_2D_2F_2$ in Figure B.3b for the alternative model. As above, the quantity $x_2(\overline{x_1}, p_{20})$, as defined for the first model, may never be observed in practice, but it can be recovered using estimates of the slope of demand curve in period 2, and of the long-term effects of the short-run policy on both quantity and price.

In contrast to the welfare effect, the formula for the bias remains essentially unchanged:

$$\begin{aligned} Bias^{LR} &= \beta \int_{x_{20}}^{\overline{x_2}} \left[p_{2,NoH}(x_2) - c_2'(x_2)\right] dx_2 - \beta \int_{x_{20}}^{x_2(\overline{x_1},p_{20})} \left[p_{20} - c_2'(x_2)\right] dx_2 \\ &- \beta \int_{x_2(\overline{x_1},p_{20})}^{\overline{x_2}} \left[p_2(x_2|\overline{x_1}, I_1(\overline{x_1})) - c_2'(x_2)\right] dx_2 - \int_{I_1(\overline{x_1})}^{I_1(\overline{x_1},\overline{x_2})} \frac{dV(\overline{x_1},\overline{x_2},I_1,I_2)}{dI_1} dI_1 \\ &= \beta \left[\int_{x_{20}}^{\overline{x_2}} \left[p_{2,NoH}(x_2) - p_{20}\right] dx_2 - \int_{x_2(\overline{x_1},p_{20})}^{\overline{x_2}} \left[p_2(x_2|\overline{x_1}, I_1(\overline{x_1})) - p_{20}\right] dx_2\right] - \int_{I_1(\overline{x_1})}^{I_1(\overline{x_1},\overline{x_2})} \frac{dV(\overline{x_1},\overline{x_2},I_1,I_2)}{dI_1} dI_1 \end{aligned} \tag{48}$$

The first source of bias still corresponds to the trapezoid $D_2B_2C_2E_2$ (for both models) because the researcher would correctly account for the loss in producer surplus whether she assumes away the possibility of hysteresis or not. Moreover, researchers making different assumptions could still use the same price variation to estimate the respective demand curve in period 2. Therefore, the key statistic to evaluate the first source of bias from assuming away hysteresis remains the difference between the baseline quantity x_{20} and the long-term effect of the short-run policy. The only qualifier is that, in the case of the first model, we would need to estimate the long-term effect of the short-run policy that we would observe in the absence of price adjustment, $x_2(\overline{x_1}, p_{20})$.

B.6.4 Considering policies that include "social incentives"

A growing literature shows that economic agents sometimes act out of social or normative motives and that these motives can be affected by changes in the agents' environment. Several papers have thus explored the role and impact of social incentives (e.g., peer pressure, social recognition, social comparison) as instruments of corrective policies (e.g., Allcott and Mullainathan, 2010). We show here that the welfare implications of hysteresis are unchanged if corrective policies include social incentives, as long as any persistent effect of a short-run policy still results from choices made in response to that policy. In contrast, we show that the welfare implications of hysteresis can be very different in the event that such a policy changed social norms persistently, creating incentives to behave in a certain way even after it ended. This case is important to consider because studies that estimate the persistent effect of social incentives rarely assess whether the incentives were themselves persistent, i.e., whether they led to persistent changes in social norms.

A. Including social incentives

Let's first maintain the assumptions of the benchmark framework but assume that the corrective policies implement their restrictions on electricity use, at least partly through a social incentive component, which we model in a similar way as, e.g., Allcott and Kessler (forthcoming). The policies add a new term in the household utility function, $\tau_t [N_t - x_t]$ with social price τ_t and social norm N_t, in the periods t in which they are in place. The norm captures an ideal electricity use level and the social price captures a social marginal incentive to reduce electricity use, e.g., a feeling of shame per unit consumed above the norm or a feeling of pride per unit reduced below the norm. For instance, τ_t corresponds to a social corrective *tax* (resp. *subsidy*) if $N_t \leq 0$ (resp. $N_t \geq x_{t0}$): it creates a marginal incentive to reduce electricity use and necessarily reduces the household's utility (resp. increases the household's utility). This modelization encompasses intermediate situations for $N_t \in (0, x_{t0})$, in which the social incentive can act first as a tax and then as a subsidy as the household reduces its electricity use x_t.

Short-run policy. Consider first the short-run policy ($\overline{x_1} < x_{10}$). Assuming away hysteresis, the researcher can still recover the change in welfare by tracing it along the path from x_{10} to $\overline{x_1}$, letting I_1 adjust endogenously, and assuming that variables in period 2 remain unaffected:

$$\begin{aligned}\Delta W^{SR}_{NoH} &= \int_{x_{10}}^{\overline{x_1}} \frac{dU_1(x_1, I_1)}{dx_1} dx_1 + \tau_1 [N_1 - \overline{x_1}] - \int_{x_{10}}^{\overline{x_1}} \frac{dE_1(x_1)}{dx_1} dx_1 \\ &= \int_{x_{10}}^{\overline{x_1}} [p_{1,NoH}(x_1) - p_{10}]\, dx_1 + \tau_1 [N_1 - \overline{x_1}] - \int_{x_{10}}^{\overline{x_1}} MD_1(x_1) dx_1 \end{aligned} \quad (49)$$

The welfare effect of the policy includes an additional term, which corresponds to the direct effect of the social incentives on the household utility (besides any change in behavior). This term appears in equation (49) because it does not constitute a transfer between the household and the government, in contrast to standard Pigouvian taxes and subsidies. Moreover, it could increase or decrease the welfare effect of the policy. For instance, a policy could increase welfare even if externalities are limited, if it creates a sense of social pride for behaving in accordance with the new norm. In practice, however, quantifying the utility gains or losses from social incentives is challenging. It would require, e.g., to elicit households' willingness to pay for a policy affecting all households, and not just the respondent herself as in Allcott and Kessler (forthcoming).

Allowing for hysteresis, the researcher can still recover the change in welfare by tracing it along the path from x_{10} to $\overline{x_1}$, letting I_1, x_2, and I_2 adjust endogenously:

$$\begin{aligned}\Delta W^{SR} &= \int_{x_{10}}^{\overline{x_1}} \frac{dU_1(x_1, I_1)}{dx_1} dx_1 + \tau_1 [N_1 - \overline{x_1}] - \int_{x_{10}}^{\overline{x_1}} \frac{dE_1(x_1)}{dx_1} dx_1 - \beta \int_{x_{10}}^{\overline{x_1}} \frac{dE_2(x_2)}{dx_2} \frac{dx_2}{dx_1} dx_1 \\ &= \int_{x_{10}}^{\overline{x_1}} [p_{1,NoH}(x_1) - p_{10}]\, dx_1 + \tau_1 [N_1 - \overline{x_1}] - \int_{x_{10}}^{\overline{x_1}} MD_1(x_1) dx_1 - \beta \int_{x_{20}}^{x_2(\overline{x_1})} MD_2(x_2) dx_2 \end{aligned} \quad (50)$$

The only difference with the benchmark framework is again the direct effect of the social incentives on the household utility. Moreover, as this term appears in the welfare formulas whether one assumes away hysteresis or not, the welfare implications of hysteresis and the bias from assuming away hysteresis remain the same as in the benchmark framework.

Long-run policy. The same conclusion holds for the long-run policy ($\overline{x_1} < x_{10}, \overline{x_2} < x_{20}$). The welfare formulas simply include an additional term for the direct effect of the social incentives, whether one assumes away hysteresis or not. Specifically, following the same approaches as in the benchmark framework, we obtain:

$$\Delta W^{LR}_{NoH} = \int_{x_{10}}^{\overline{x_1}} [p_{1,NoH}(x_1) - p_{10}]\, dx_1 + \beta \int_{x_{20}}^{\overline{x_2}} [p_{2,NoH}(x_2) - p_{20}]\, dx_2 + \tau_1 [N_1 - \overline{x_1}] + \beta \tau_2 [N_2 - \overline{x_2}] - \int_{x_{10}}^{\overline{x_1}} MD_1(x_1) dx_1 - \beta \int_{x_{20}}^{\overline{x_2}} MD_2(x_2) dx_2 \tag{51}$$

$$\Delta W^{LR} = \int_{x_{10}}^{\overline{x_1}} [p_1(x_1) - p_{10}]\, dx_1 + \beta \int_{x_2(\overline{x_1})}^{\overline{x_2}} [p_2(x_2|\overline{x_1}, I_1(\overline{x_1})) - p_{20}]\, dx_2 + \int_{I_1(\overline{x_1})}^{I_1(\overline{x_1},\overline{x_2})} \frac{dV(\overline{x_1},\overline{x_2},I_1,I_2)}{dI_1} dI_1 + \tau_1 [N_1 - \overline{x_1}] + \beta \tau_2 [N_2 - \overline{x_2}] - \int_{x_{10}}^{\overline{x_1}} MD_1(x_1) dx_1 - \beta \int_{x_{20}}^{\overline{x_2}} MD_2(x_2) dx_2 \tag{52}$$

So, the welfare implications of hysteresis and the bias from assuming away hysteresis remain the same as in the benchmark framework.

B. Including social incentives and assuming that the persistent effect of the short-run policy is due to persistent changes in social norms.

As above, let's assume that the policies implement their restrictions on electricity use, at least partly through a social incentive component. However, let's assume now that the propensity to consume does not in fact depend on past choices – i.e., we have $s_t(s_0, I_t)$ – but that the short-run policy creates a persistent change in social norms in period 2, captured by a social incentive $\tau_{\bar{x}_1} [N_{\bar{x}_1} - x_2]$, even though the policy itself ended. This is the only reason why the household uses less electricity in period 2 following the short-run policy.

Short-run policy. A researcher assuming away hysteresis will assume away the persistence of any social incentive. Therefore, she will derive the same welfare effect for the short-run policy ($\overline{x_1} < x_{10}$) as in equation (49). In contrast, the welfare effect differs when we take into account the persistence of the social incentive. Following the usual approach, we obtain:

$$\Delta W^{SR} = \int_{x_{10}}^{\overline{x_1}} [p_1(x_1) - p_{10}]\, dx_1 + \beta \int_{x_{20}}^{x_2(\overline{x_1})} [p_2(x_2) - p_{20}]\, dx_2 + \tau_1 [N_1 - \overline{x_1}] + \beta \tau_{\bar{x}_1} [N_{\bar{x}_1} - x_2] - \int_{x_{10}}^{\overline{x_1}} MD_1(x_1) dx_1 - \beta \int_{x_{20}}^{x_2(\overline{x_1})} MD_2(x_2) dx_2 \tag{53}$$

The change in private welfare includes four terms in equation (53). The first integral is the usual loss in private welfare from inducing the household to consume below its privately optimal level in period 1. The second integral now also captures the loss in private welfare from the persistent social incentive inducing the household to consume below its privately optimal level in period 2. This is not captured by the demand curve in period 1 because the household's choices in period 1 do not affect its choices in period 2 in this model. Instead, it can be captured by a standard Harberger triangle under the demand curve in period 2, which is not indexed by $\overline{x_1}$ anymore. The long-run effect of the short-run policy remains a key empirical object because, on top of being key to evaluate the correction of the externality in period 2, it forms the base of that triangle. Finally, the third and fourth terms capture the direct gain or loss in private welfare resulting from the social incentive included in the short-run policy and from the persistent incentive in period 2, respectively.

The sign of the bias from assuming away hysteresis is now ambiguous:

$$Bias^{SR} = \beta \int_{x_{20}}^{x_2(\overline{x_1})} MD_2(x_2) dx_2 - \beta \int_{x_{20}}^{x_2(\overline{x_1})} [p_2(x_2) - p_{20}] \, dx_2 - \beta \tau_{\bar{x}_1} [N_{\bar{x}_1} - x_2] \tag{54}$$

The bias depends on the relative sizes of the welfare gain from the correction of the externality in period 2, of the welfare loss from the Harberger triangle in period 2, and of the direct effect of the persistent social incentive on the household utility. The long-run effect of the short-run policy is a key statistic to evaluate the first two terms. The third term is challenging to estimate and can again be positive or negative, depending on whether the social incentive acts as a social tax or subsidy.

Long-run policy. A researcher assuming away hysteresis will again assume away the persistence of the social incentive. So, she will derive the same welfare effect for the long-run policy ($\overline{x_1} < x_{10}, \overline{x_2} < x_{20}$) as in equation (51). The welfare formula would in fact also be identical if we were to take into account the persistence of the social incentive. There is no other mechanism of hysteresis, so none of the three sources of bias in the benchmark framework arises. As a result, there is no bias from assuming away hysteresis in this model for the case of the long-run policy.[75]

This extension highlights again the importance of studying *why* agents make persistent changes to understand the welfare implications of hysteresis. For instance, in our own application, we argue that persistent changes in social norms are unlikely to rationalize some of our findings.

B.6.5 Allowing for present bias or myopic beliefs

The slopes of the demand curves are sufficient statistics to evaluate changes in private welfare in the benchmark framework because the household maximizes the same private welfare that is

[75]Note that we are implicitly assuming that the social incentive created by the long-run policy in period 2, $\tau_2 [N_2 - \overline{x_2}]$, is the binding one in period 2, and thus not $\tau_{\bar{x}_1} [N_{\bar{x}_1} - x_2]$. This makes sense if the household only reduces electricity use to $x_2(\overline{x_1})$ with the incentive $\tau_{\bar{x}_1} [N_{\bar{x}_1} - x_2]$, but to $\overline{x_2} < x_2(\overline{x_1})$ with the incentive $\tau_2 [N_2 - \overline{x_2}]$. The fact that some social incentive would exist in period 2 even in absence of corrective policy in that period, may further complicates any attempt at estimating the direct effect of the social incentive $\tau_2 [N_2 - \overline{x_2}]$.

included in the social welfare function. This rests on relatively standard assumptions, for instance that households have time-consistent preferences and that they are fully aware of the effect of their current choices on their future propensity to consume. However, some of these assumptions may be relatively strong in some contexts, including in the context of our empirical application. We thus show here how relaxing some of these assumptions affects our main takeaways. We focus on two models that constitute first natural deviations from the benchmark framework: (i) a model in which the representative household is present-biased, and (ii) a model in which the household has myopic beliefs about the effect of its current choices on its future propensity to consume.

A. Present bias

We start with the case of a present-biased household. We assume that the household makes decisions in period 1 by maximizing its perceived lifetime utility $\widetilde{V}$, discounting the future by $\alpha\beta$ instead of β, where $\alpha \in [0,1)$ captures a bias towards the present:

$$\max_{x_1,x_2,I_1,I_2} \widetilde{V} = U_1 + \alpha\beta U_2 = y_1 - p_1 x_1 - \kappa_1(I_1) + v_1(x_1,s_1) + \alpha\beta\left[y_2 - p_2 x_2 - \kappa_2(I_2) + v_2(x_2,s_2)\right]$$

$$\text{s.t. } s_t = s_t(s_{t-1}, x_{t-1}, I_t) \text{ for } t = 1,2 \text{ and } s_0, x_0 \text{ given} \tag{55}$$

In this model, the first-order conditions for the household problem are:

$$\frac{\partial v_1(x_1,s_1)}{\partial x_1} + \alpha\beta \frac{\partial v_2(x_2,s_2)}{\partial s_2}\frac{\partial s_2}{\partial x_1} = p_1 \quad ; \quad \frac{\partial v_2(x_2,s_2)}{\partial x_2} = p_2 \tag{56}$$

$$\left[\frac{\partial v_1(x_1,s_1)}{\partial s_1} + \alpha\beta \frac{\partial v_2(x_2,s_2)}{\partial s_2}\frac{\partial s_2}{\partial s_1}\right]\frac{\partial s_1}{\partial I_1} - \kappa_1'(I_1) = 0 \quad ; \quad \frac{\partial v_2(x_2,s_2)}{\partial s_2}\frac{\partial s_2}{\partial I_2} - \kappa_2'(I_2) = 0 \tag{57}$$

Equations (56) and (57) imply that a present-biased household uses more electricity in period 1 and invests less in reducing its propensity to consume. Conditional on period-1 choices, however, the household makes privately optimal choices in period 2, that it planned to make while in period 1.[76]

There is now a potential disagreement between the household and the government regarding the relevant notion of private welfare, depending on the welfare criterion adopted by the government. On the one hand, the government could evaluate private welfare based on the household's "experienced" utility, discounting the future without a bias for the present ($\alpha = 1$). This is the welfare criterion suggested in, e.g., Farhi and Gabaix (2017). In this case, the private welfare simply corresponds to the household lifetime utility, V, in equation (16) and welfare is

[76] In a model with three periods, the household consumes too much electricity and underinvests in its propensity to consume in period 2, even from its own period-1 perspective. As a result, the corrective policies further increase welfare under both welfare criteria by correcting an internality in period 2 in the right direction (see below).

$W^e = V - E_1(x_1) - \beta E_2(x_2)$, where e stands for "experienced." This is also equivalent to evaluating private welfare based on some "period-0" utility.[77] On the other hand, the government could evaluate private welfare based on the household's "decision" utility from the perspective of period 1. In other words, it could evaluate private welfare using equation (55). In this case, welfare is: $W^d = \widetilde{V} - E_1(x_1) - \alpha\beta E_2(x_2)$, where d stands for "decision".[78]

It is straightforward that all the results in the benchmark framework carry through with the decision utility criterion, as the household maximizes again the same private welfare that is included in the social welfare function. The terms related to period 2 must simply be scaled down by α.

In contrast, results are different from those in the benchmark framework with the experienced utility criterion. To see this, it is useful to rewrite welfare as: $W^e = \widetilde{V} + (V - \widetilde{V}) - E_1(x_1) - \beta E_2(x_2)$, where the difference between decision and experienced utility $(V - \widetilde{V})$ captures an *internality*. Corrective policies can generate a new source of welfare gain by addressing this internality, moving consumption and investment choices towards their true private optimum. We show this below where we use the experienced utility criterion to evaluate the welfare effect of the short- and long-run policies, given the choices made by the household according to equations (56) and (57).

Short-run policy. Let's start with the short-run policy ($\overline{x_1} < x_{10}$), and let's define $\widetilde{x}_2(x_1), \widetilde{I}_1(x_1)$, and $\widetilde{I}_2(x_1)$ as the household choices according to equations (56) and (57) for a given value of x_1.

Assuming away hysteresis, the researcher assumes that choices in the two periods are independent. As a result, a bias towards the present is inconsequential; the parameter α does not appear in equations (56) and (57) if we assume $s_t(s_0, I_t)$. Moreover, the household values changes in x_1 on its private welfare accurately, as they only affect its utility in period 1. So the slope of the demand curve in period 1, which is a function of the household perceived change in its private welfare, remains a sufficient statistic for the actual change in its private welfare following changes in x_1:

$$\begin{aligned}
\Delta W^{e,SR}_{NoH} &= \int_{x_{10}}^{\overline{x_1}} \frac{d\widetilde{V}(x_1,x_2,I_1,I_2)}{dx_1}dx_1 + \int_{x_{10}}^{\overline{x_1}} \left[\frac{dV(x_1,\widetilde{x}_2(x_1),\widetilde{I}_1(x_1),\widetilde{I}_2(x_1))}{dx_1} - \frac{d\widetilde{V}(x_1,x_2,I_1,I_2)}{dx_1}\right] dx_1 - \int_{x_{10}}^{\overline{x_1}} \frac{dE_1(x_1)}{dx_1}dx_1 \\
&= \int_{x_{10}}^{\overline{x_1}} \frac{dU_1(x_1,I_1)}{dx_1}dx_1 + \int_{x_{10}}^{\overline{x_1}} \left[\frac{dU_1(x_1,I_1)}{dx_1} - \frac{dU_1(x_1,I_1)}{dx_1}\right] dx_1 - \int_{x_{10}}^{\overline{x_1}} \frac{dE_1(x_1)}{dx_1}dx_1 \\
&= \int_{x_{10}}^{\overline{x_1}} [p_{1,NoH}(x_1) - p_{10}]\, dx_1 - \int_{x_{10}}^{\overline{x_1}} MD_1(x_1)dx_1 \qquad (58)
\end{aligned}$$

In contrast, the researcher must account for differences between the private welfare as maxi-

[77] In that case, the private welfare would be evaluated at $V_0 = \alpha\beta V$; the preference for the present would be irrelevant because no choices are made before period 1; and welfare would be $W^0 = \alpha\beta\left[V - E_1(x_1) - \beta E_2(x_2)\right]$.

[78] If it makes sense to evaluate private welfare from the period-1 perspective, it makes sense to apply the same approach to evaluate the welfare consequences of the externalities, although this is not important for our results.

mized by the household and as evaluated by the government when allowing for hysteresis:

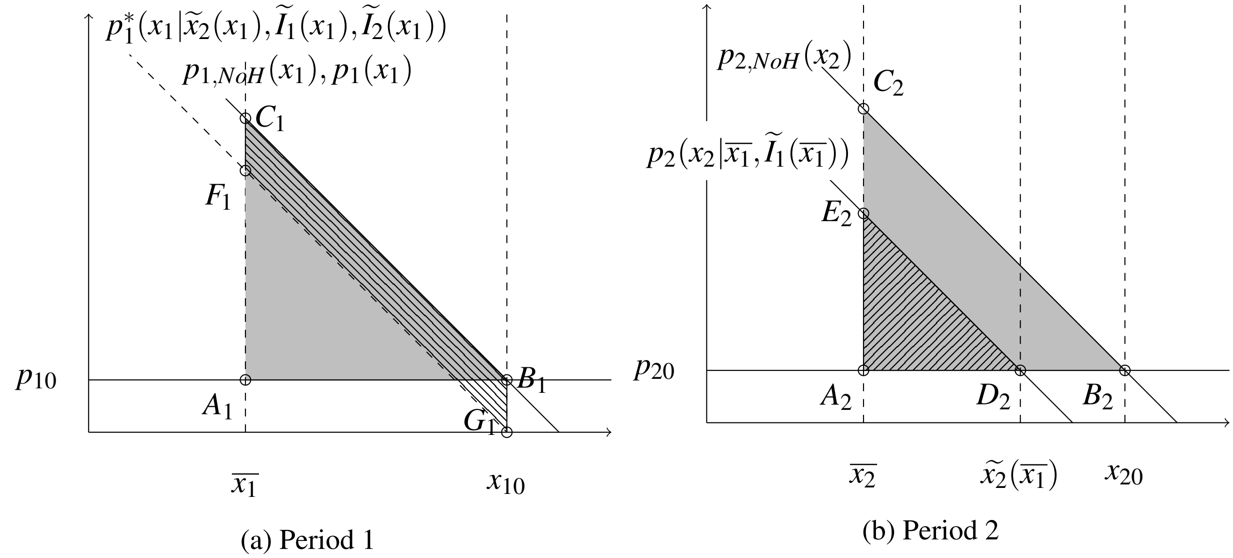
$$\Delta W^{e,SR} = \int_{x_{10}}^{\overline{x_1}} \frac{d\widetilde{V}(x_1,x_2,I_1,I_2)}{dx_1}dx_1 + \int_{x_{10}}^{\overline{x_1}} \left[\frac{dV(x_1,\widetilde{x}_2(x_1),\widetilde{I}_1(x_1),\widetilde{I}_2(x_1))}{dx_1} - \frac{d\widetilde{V}(x_1,x_2,I_1,I_2)}{dx_1} \right] dx_1$$
$$- \int_{x_{10}}^{\overline{x_1}} \frac{dE_1(x_1)}{dx_1}dx_1 - \beta \int_{x_{20}}^{\widetilde{x}_2(\overline{x_1})} \frac{dE_2(x)}{dx_2}dx_2$$
$$= \int_{x_{10}}^{\overline{x_1}} [p_1(x_1) - p_{10}]\, dx_1 + \int_{x_{10}}^{\overline{x_1}} \left[p_1^*(x_1|\widetilde{x}_2(x_1),\widetilde{I}_1(x_1),\widetilde{I}_2(x_1)) - p_1(x_1) \right] dx_1$$
$$- \int_{x_{10}}^{\overline{x_1}} MD_1(x_1)dx_1 - \beta \int_{x_{20}}^{\widetilde{x}_2(\overline{x_1})} MD_2(x_2)dx_2 \tag{59}$$

The new term in equation (59) corresponds to the *average marginal bias* (Allcott and Taubinsky, 2015) between $\overline{x_1}$ and x_{10}. It is the difference between the actual utility derived from a given unit of x_1 (besides its monetary cost), which is captured by the function $p_1^*(x_1|\widetilde{x}_2(x_1),\widetilde{I}_1(x_1),\widetilde{I}_2(x_1))$, and the household perceived utility from a given unit of x_1, which is captured by the demand curve. For instance, the household systematically overestimates its utility gain from a marginal unit of x_1 in our model because of the present bias. The function p_1^* would thus lie below the demand curve, as depicted in Figure B.4a. The average marginal bias corresponds to the trapezoid $B_1C_1F_1G_1$ in that case. Our notation emphasizes that p_1^* captures the actual utility derived from x_1 given the household's decision regarding the levels of the other choice variables, which may not be optimal. Therefore, this empirical object is not easily observed in practice. In particular, it does not correspond to the demand curve that would be observed if the household was de-biased. Indeed, in that case, the change in the demand curve would not capture the average marginal bias, as the new demand curve will not reflect the household's accurate valuation of changes in x_1 holding constant its choice of the other variables for a given value of x_1. This is because its choices of x_2, I_1, and I_2 for a given value of x_1 would likely change as well. The demand curve that would be observed if the household was de-biased would thus also reflect the household new valuation of changes in x_1 given its new choices for the other variables. This complication does not arise in Allcott and Taubinsky (2015) because their model includes a single choice variable.

Importantly, the average marginal bias can be signed in many cases even if it cannot be estimated directly. For instance, it is positive in our case because the policy corrects the internality in the right direction. As a result, the welfare effect of the short-run policy and the bias from assuming away hysteresis, in which case the researcher assumes away the internality issue, would increase with the experienced utility criterion in this model compared to the benchmark framework.

Long-run policy. Let's consider now the long-run policy ($\overline{x_1} < x_{10}$, $\overline{x_2} < x_{20}$). Assuming away hysteresis, choices in periods 1 and 2 are assumed to be independent, so the bias is again inconse-

Figure B.4: The loss in private welfare from corrective policies with and without hysteresis in a model with present bias



The figure illustrates the loss in private welfare in a model with a present-biased household (the private welfare corresponds to the consumer surplus) for the short-run and the long-run corrective polices. For simplicity, we assume linear demand curves and no change in the slope of the demand curve following a short-run policy in the presence of hysteresis. See text for details.

quential, and we obtain the same formula for the welfare effect as in the benchmark framework:

$$\begin{aligned}\Delta W_{NoH}^{e,LR} &= \int_{x_{10}}^{\overline{x_1}} \frac{d\widetilde{V}(x_1,x_2,I_1,I_2)}{dx_1}dx_1 + \int_{x_{10}}^{\overline{x_1}} \left[\frac{dV(x_1,\widetilde{x_2}(x_1),\widetilde{I_1}(x_1),\widetilde{I_2}(x_1))}{dx_1} - \frac{d\widetilde{V}(x_1,x_2,I_1,I_2)}{dx_1}\right] dx_1 \\ &+\beta \int_{x_{20}}^{\overline{x_2}} \frac{dU_2(x_2,I_2)}{dx_2}dx_2 - \int_{x_{10}}^{\overline{x_1}} \frac{dE_1(x_1)}{dx_1}dx_1 - \beta \int_{x_{20}}^{\overline{x_2}} \frac{dE_2(x_2)}{dx_2}dx_2 \qquad (60) \\ &= \int_{x_{10}}^{\overline{x_1}} [p_{1,NoH}(x_1) - p_{10}]\, dx_1 - \int_{x_{10}}^{\overline{x_1}} MD_1(x_1)dx_1 + \beta \int_{x_{20}}^{\overline{x_2}} [p_{2,NoH}(x_2) - p_{20}]\, dx_2 - \beta \int_{x_{20}}^{\overline{x_2}} MD_2(x_2)dx_2\end{aligned}$$

It is useful to consider the same three steps as before when allowing for hysteresis. The first step captures the effect of the policy in period 1, as if there was only a short-run policy. For this step, let's define again $\widetilde{x_2}(x_1)$, $\widetilde{I_1}(x_1)$, and $\widetilde{I_2}(x_1)$ as the choices according to equations (56) and (57) for a given value of x_1. The second step captures the effect of the policy in period 2, holding fixed investment levels from period 1. For this step, let's define $\widetilde{I_2}(\overline{x_1},\overline{x_2},\widetilde{I_1}(\overline{x_1}))$ as the resulting investment choices in period 2. Finally, the third step captures the fact that the household could have anticipated the policy in period 2. For this step, let's define $\widetilde{I_1}(\overline{x_1},\overline{x_2})$ and $\widetilde{I_2}(\overline{x_1},\overline{x_2})$, the actual investment choices in the two periods given the long-run policy. The welfare effect is then:

$$\begin{aligned}
\Delta W^{e,LR} = & \int_{x_{10}}^{\overline{x_1}} \frac{d\widetilde{V}(x_1,x_2,I_1,I_2)}{dx_1}dx_1 + \int_{x_{10}}^{\overline{x_1}} \left[\frac{dV(x_1,\widetilde{x}_2(x_1),\widetilde{I}_1(x_1),\widetilde{I}_2(x_1))}{dx_1} - \frac{d\widetilde{V}(x_1,x_2,I_1,I_2)}{dx_1}\right] dx_1 \\
& + \beta \int_{\widetilde{x}_2(\overline{x}_1)}^{\overline{x_2}} \frac{dU_2(\overline{x}_1,x_2,\widetilde{I}_1(\overline{x}_1),I_2)}{dx_2}dx_2 + \int_{\widetilde{I}_1(\overline{x_1})}^{\widetilde{I}_1(\overline{x_1},\overline{x_2})} \frac{d\widetilde{V}(\overline{x_1},\overline{x_2},I_1,I_2)}{dI_1}dI_1 \\
& - \int_{x_{10}}^{\overline{x_1}} \frac{dE_1(x_1)}{dx_1}dx_1 - \beta \int_{x_{20}}^{\overline{x_2}} \frac{dE_2(x_2)}{dx_2}dx_2 \\
= & \int_{x_{10}}^{\overline{x_1}} [p_1(x_1) - p_{10}]\,dx_1 + \int_{x_{10}}^{\overline{x_1}} \left[p_1^*(x_1|\widetilde{x}_2(x_1),\widetilde{I}_1(x_1),\widetilde{I}_2(x_1)) - p_1(x_1)\right] dx_1 \\
& + \beta \int_{\widetilde{x}_2(\overline{x}_1)}^{\overline{x_2}} \left[p_2(x_2|\overline{x_1},\widetilde{I}_1(\overline{x_1})) - p_{20}\right] dx_2 + \int_{\widetilde{I}_1(\overline{x_1})}^{\widetilde{I}_1(\overline{x_1},\overline{x_2})} \frac{d\widetilde{V}(\overline{x_1},\overline{x_2},I_1,I_2)}{dI_1}dI_1 \\
& - \int_{x_{10}}^{\overline{x_1}} MD_1(x_1)dx_1 - \beta \int_{x_{20}}^{\overline{x_2}} MD_2(x_2)dx_2
\end{aligned} \tag{61}$$

The only novelty in equation (61) comes again from the correction of the internality in period 1. In particular, there is no internality issue in period 2 given period-1 choices, so the slope of the demand curve in period 2 remains a sufficient statistic for the change in private welfare from the change in x_2 caused by the period-2 policy.[79] As a result, there is no difference for period 2 in Figure B.4b. Moreover, the change in welfare from taking into account that the household could have anticipated the period-2 policy in period 1 remains (weakly) positive. A present-biased household would underinvest in period 1, so the fact that the household increases its period-1 investments in anticipation of the period-2 policy can only increase welfare. Additionally, the fact that the household may also change its period-2 investments cannot decrease welfare because the household makes optimal choices in period 2 given the choices it made in period 1.

In sum, our main takeaway is only strengthened in this model when internality and externality issues are aligned. The formulas simply add a term for the average marginal bias in period 1.

B. Myopic household

We now consider the case of a household who is myopic about the effect of its current choices on its future propensity to consume. We assume that the household maximizes its perceived lifetime utility in period 1, assuming away any effect of its choices on its future propensity to consume, but that it makes decisions in period 2 based on its actual propensity to consume.

[79] As mentioned above, in a model with three periods, the household consumes too much electricity and underinvests in its propensity to consume in period 2, even from its own period-1 perspective. As a result, the corrective policies would further increase welfare under both criteria by correcting an internality in period 2 in the right direction.

Formally, the household solves the following problem in period 1:

$$\max_{x_1,x_2,I_1,I_2} \widetilde{V} = U_1 + \beta U_2(s_0) = y_1 - p_1 x_1 - \kappa_1(I_1) + \nu_1(x_1,s_1) + \beta\left[y_2 - p_2 x_2 - \kappa_2(I_2) + \nu_2(x_2,s_2)\right]$$
$$\text{s.t. } s_1 = s_1(s_0,I_1) \text{ and } s_2 = s_2(s_0,I_2) \text{ for some } s_0 \text{ given} \tag{62}$$

It thus makes its choices in period 1, and expects to make choices in period 2, according to the following first-order conditions:

$$\frac{\partial \nu_1(x_1,s_1)}{\partial x_1} = p_1 \quad ; \quad \frac{\partial \nu_2(x_2,s_2(s_0,x_0,I_2))}{\partial x_2} = p_2 \tag{63}$$

$$\frac{\partial \nu_1(x_1,s_1)}{\partial s_1}\frac{\partial s_1}{\partial I_1} - \kappa_1'(I_1) = 0 \quad ; \quad \frac{\partial \nu_2(x_2,s_2(s_0,x_0,I_2))}{\partial s_2}\frac{\partial s_2}{\partial I_2} - \kappa_2'(I_2) = 0 \tag{64}$$

However, the household solves a different problem in period 2 than it anticipated in period 1:

$$\max_{x_2,I_2} U_2 = y_2 - p_2 x_2 - \kappa_2(I_2) + \nu_2(x_2,s_2(s_1,x_1,I_2)) \tag{65}$$

It thus actually makes choices in period 2 according to the first-order conditions below:

$$\frac{\partial \nu_2(x_2,s_2(s_1,x_1,I_2))}{\partial x_2} = p_2 \quad ; \quad \frac{\partial \nu_2(x_2,s_2(s_1,x_1,I_2))}{\partial s_2}\frac{\partial s_2}{\partial I_2} - \kappa_2'(I_2) = 0 \tag{66}$$

Such a myopic household consumes more electricity and invests less in reducing its propensity to consume in period 1. Moreover, by assuming away hysteresis, it fails to predict its propensity to consume accurately in period 2, and thus its choices in that period.

There is now a potential disagreement between the household and the government regarding the relevant notion of private welfare, depending on the welfare criterion adopted by the government. On the one hand, the government could evaluate private welfare based on the household's "experienced" utility. The private welfare would then correspond to the household lifetime utility, V, in equation (16) and welfare would be: $W^e = \widetilde{V} + (V - \widetilde{V}) - E_1(x_1) - \beta E_2(x_2)$. On the other hand, the government could evaluate private welfare based on the household's "decision" utility in period 1, using the household expected utility function, $\widetilde{V}$, in equation (62). In other words, the government could evaluate private welfare abstracting from the changes in utility associated with the underlying mechanisms of hysteresis. This could make sense in some cases, e.g., if the impact of its choices on its future utility is actually seen as harmful for the household, as with addictions. In this case, welfare is: $W^d = \widetilde{V} - E_1(x_1) - \beta E_2(x_2)$. In both cases, the disagreement between the household and the government will create a second potential source of welfare gain or loss from corrective policies. In the first case, the individual and the government will disagree on the privately optimal behavior in period 1. In the second case, they will disagree on the privately optimal behavior in period 2. We thus consider both cases below.

Short-run policy. We start with the short-run policy ($\overline{x_1} < x_{10}$). Let's define $\widetilde{I}_1(x_1)$ as the investment choices in period 1 according to equations (63) and (64) for a given value of x_1. Let's also define $\widetilde{x}_2$ and $\widetilde{I}_2$ as the expected choices for period 2, which the household assumes to be independent of x_1 and I_1 in equations (63) and (64). Finally, let's define $x_2(x_1, \widetilde{I}_1(x_1))$ and $I_2(x_1, x_2, \widetilde{I}_1(x_1))$ as the actual choices in period 2 given the first-order conditions (66).

Assuming away hysteresis, decisions in the two periods are assumed to be independent. There is thus no room for myopia to matter; the household's "decision" utility and "experienced" utility are assumed to coincide, and the welfare formula is the same with both criteria:

$$\begin{aligned}
\Delta W^{e,SR}_{NoH} &= \int_{x_{10}}^{\overline{x_1}} \frac{d\widetilde{V}(x_1,x_2,I_1,I_2)}{dx_1} dx_1 + \int_{x_{10}}^{\overline{x_1}} \left[\frac{dV(x_1,\widetilde{x}_2,\widetilde{I}_1(x_1),\widetilde{I}_2)}{dx_1} - \frac{d\widetilde{V}(x_1,x_2,I_1,I_2)}{dx_1} \right] dx_1 - \int_{x_{10}}^{\overline{x_1}} \frac{dE_1(x_1)}{dx_1} dx_1 \\
&= \int_{x_{10}}^{\overline{x_1}} \frac{dU_1(x_1,I_1)}{dx_1} dx_1 + \int_{x_{10}}^{\overline{x_1}} \left[\frac{dU_1(x_1,I_1)}{dx_1} - \frac{dU_1(x_1,I_1)}{dx_1} \right] dx_1 - \int_{x_{10}}^{\overline{x_1}} \frac{dE_1(x_1)}{dx_1} dx_1 \\
&= \int_{x_{10}}^{\overline{x_1}} \left[p_{1,NoH}(x_1) - p_{10} \right] dx_1 - \int_{x_{10}}^{\overline{x_1}} MD_1(x_1) dx_1 \qquad (67)
\end{aligned}$$

$$\begin{aligned}
\Delta W^{d,SR}_{NoH} &= \int_{x_{10}}^{\overline{x_1}} \frac{d\widetilde{V}(x_1,x_2,I_1,I_2)}{dx_1} dx_1 - \int_{x_{10}}^{\overline{x_1}} \frac{dE_1(x_1)}{dx_1} dx_1 = \int_{x_{10}}^{\overline{x_1}} \frac{dU_1(x_1,I_1)}{dx_1} dx_1 - \int_{x_{10}}^{\overline{x_1}} \frac{dE_1(x_1)}{dx_1} dx_1 \\
&= \int_{x_{10}}^{\overline{x_1}} \left[p_{1,NoH}(x_1) - p_{10} \right] dx_1 - \int_{x_{10}}^{\overline{x_1}} MD_1(x_1) dx_1 \qquad (68)
\end{aligned}$$

In contrast, when allowing for hysteresis, the researcher must account for differences between the private welfare as maximized by the household and as evaluated by the government:

$$
\begin{aligned}
\Delta W^{e,SR} =& \int_{x_{10}}^{\overline{x_1}} \frac{d\widetilde{V}(x_1,x_2,I_1,I_2)}{dx_1}dx_1 + \int_{x_{10}}^{\overline{x_1}} \left[\frac{dV(x_1,\widetilde{x}_2,\widetilde{I}_1(x_1),\widetilde{I}_2)}{dx_1} - \frac{d\widetilde{V}(x_1,x_2,I_1,I_2)}{dx_1}\right] dx_1 \\
&+ \beta \int_{\widetilde{x}_2}^{x_2(\overline{x_1},\widetilde{I}_1(\overline{x_1}))} \frac{dU_2(\overline{x_1},x_2,\widetilde{I}_1(\overline{x_1}),I_2)}{dx_2}dx_2 \\
&- \int_{x_{10}}^{\overline{x_1}} \frac{dE_1(x_1)}{dx_1}dx_1 - \beta \int_{x_{20}}^{x_2(\overline{x_1},\widetilde{I}_1(\overline{x_1}))} \frac{dE_2(x)}{dx_2}dx_2 \\
=& \int_{x_{10}}^{\overline{x_1}} [p_1(x_1) - p_{10}]\,dx_1 + \int_{x_{10}}^{\overline{x_1}} \left[p_1^*(x_1|\widetilde{x}_2,\widetilde{I}_1(x_1),\widetilde{I}_2) - p_1(x_1)\right] dx_1 \\
&+ \beta \int_{\widetilde{x}_2}^{x_2(\overline{x_1},\widetilde{I}_1(\overline{x_1}))} \left[p_2(x_2|\overline{x_1},\widetilde{I}_1(\overline{x_1})) - p_{20}\right] dx_2 \\
&- \int_{x_{10}}^{\overline{x_1}} MD_1(x_1)dx_1 - \beta \int_{x_{20}}^{x_2(\overline{x_1},\widetilde{I}_1(\overline{x_1}))} MD_2(x_2)dx_2
\end{aligned} \tag{69}
$$

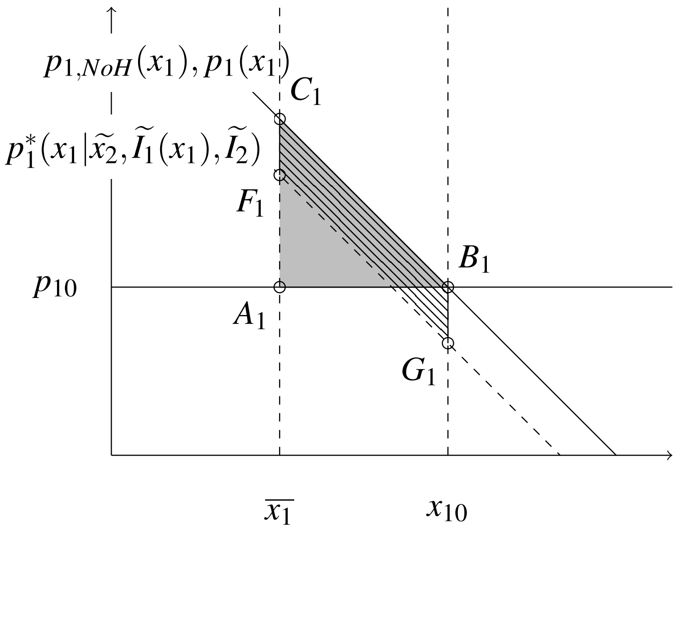
$$
\begin{aligned}
\Delta W^{d,SR} =& \int_{x_{10}}^{\overline{x_1}} \frac{d\widetilde{V}(x_1,x_2,I_1,I_2)}{dx_1}dx_1 + \beta \int_{\widetilde{x}_2}^{x_2(\overline{x_1},\widetilde{I}_1(\overline{x_1}))} \frac{dU_2(\overline{x_1},x_2,\widetilde{I}_1(\overline{x_1}),I_2)}{dx_2}dx_2 \\
&+ \beta \int_{\widetilde{x}_2}^{x_2(\overline{x_1},\widetilde{I}_1(\overline{x_1}))} \left[\frac{dU_2(s_0,x_2,I_2(\overline{x_1},x_2,\widetilde{I}_1(\overline{x_1})))}{dx_2} - \frac{dU_2(\overline{x_1},x_2,\widetilde{I}_1(\overline{x_1}),I_2)}{dx_2}\right] dx_2 \\
&- \int_{x_{10}}^{\overline{x_1}} \frac{dE_1(x_1)}{dx_1}dx_1 - \beta \int_{x_{20}}^{x_2(\overline{x_1},\widetilde{I}_1(\overline{x_1}))} \frac{dE_2(x)}{dx_2}dx_2 \\
=& \int_{x_{10}}^{\overline{x_1}} [p_1(x_1) - p_{10}]\,dx_1 + \beta \int_{\widetilde{x}_2}^{x_2(\overline{x_1},\widetilde{I}_1(\overline{x_1}))} \left[p_2(x_2|\overline{x_1},\widetilde{I}_1(\overline{x_1})) - p_{20}\right] dx_2 \\
&+ \beta \int_{\widetilde{x}_2}^{x_2(\overline{x_1},\widetilde{I}_1(\overline{x_1}))} \left[p_2^*(x_2|s_0,I_2(\overline{x_1},x_2,\widetilde{I}_1(\overline{x_1}))) - p_2(x_2|\overline{x_1},\widetilde{I}_1(\overline{x_1}))\right] dx_2 \\
&- \int_{x_{10}}^{\overline{x_1}} MD_1(x_1)dx_1 - \beta \int_{x_{20}}^{x_2(\overline{x_1},\widetilde{I}_1(\overline{x_1}))} MD_2(x_2)dx_2
\end{aligned} \tag{70}
$$

Using the experienced utility criterion, there are two additional terms compared to the benchmark framework in equation (69). First, there is the average marginal bias of changing consumption from x_{10} to $\overline{x_1}$, holding fixed the expected choices of the other variables according to equations (63) and (64). As for the model with present bias, it arises because the myopic household overestimates its utility gain from a unit of x_1 in period 1. It corresponds to the trapezoid $B_1C_1F_1G_1$ in Figure B.5a. The average marginal bias is again challenging to estimate, but it can be signed: the welfare effect of the short-run policy, and thus the bias from assuming away hysteresis, increases because the policy corrects the internality issue in the right direction. Second, the third term in equation (69) captures the change in private welfare from the household deciding in period 2 to revise the choices it planned to make while in period 1. This effect is necessarily positive, as the household doesn't have to revise its choice, further increasing the welfare effect of the short-run policy and the bias from assuming away hysteresis. Moreover, this effect can be recovered using the demand curve in period 2 following the short-run policy, provided we know $\widetilde{x}_2$. A natural assumption is that the household expects to consume in period 2 what it would have consumed in absence of the corrective policy, x_{20}. In other words, the household is assumed to be myopic regarding the effect of changes in period-1 consumption away from baseline consumption levels, but it is not systematically wrong about its future propensity to consume at baseline. In this case, the third term in equation (69) becomes a function of the slope of the demand curve in period 2 and of the long-term effect of the short-run policy, $x_{20} - x_2(\overline{x_1}, \widetilde{I}_1(\overline{x_1}))$. This is illustrated in Figure B.5b. At x_{20}, the household marginal utility from x_2 would be below the price p_{20} given the change in its propensity to consume caused by the short-run policy. The household thus increases its private welfare by reducing its consumption from x_{20} to $x_2(\overline{x_1}, \widetilde{I}_1(\overline{x_1}))$. This gain in private welfare is represented by the triangle $B_2D_2G_2$ in Figure B.5b, the area between the period-2 demand curve and the price level between consumption levels x_{20} and $x_2(\overline{x_1}, \widetilde{I}_1(\overline{x_1}))$. The long-run effect of the short-run policy is then again a key empirical object to evaluate the welfare effect of the short-run policy.[80]

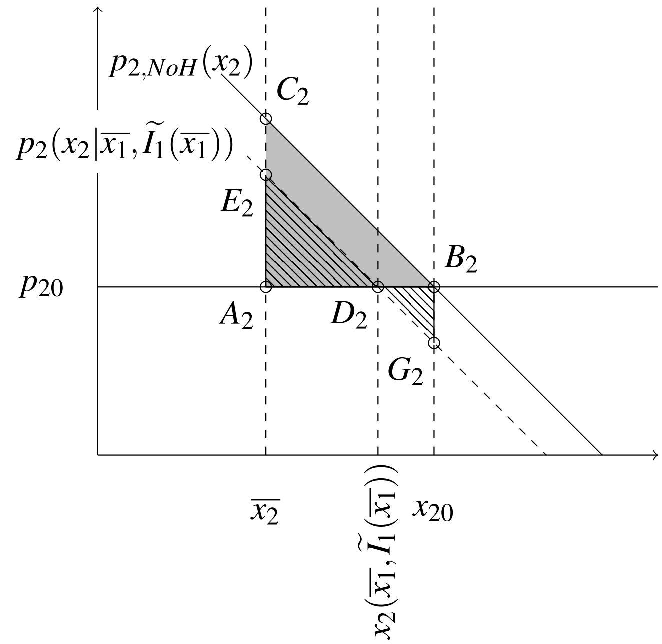
Evaluating the household private welfare based on its period-1 decision utility, the formula for the welfare effect of the short-run policy differs in important ways. First, there is no internality issue in period 1 anymore, so there is no average marginal bias in period 1 in equation (70). The loss in private welfare in period 1 is simply the usual Harberger triangle, as illustrated in Figure B.5c. The second term in equation (70) captures the same perceived increase in private welfare from the household deciding in period 2 to revise its choices as in equation (69). This is illustrated by the triangle $B_2D_2G_2$ in Figure B.5d. However, evaluating private welfare based on the household period-1 decision utility, the change in consumption levels in period 2 is considered to be a mistake. This leads to an average marginal bias in period 2, which is captured by the third term in equation (70), the difference between the actual and perceived marginal utility for units of x_2, holding fixed the household choices regarding the levels of the other variables, between expected and actual consumption levels in period 2. The function p_2^* is difficult to estimate in practice, but it must lie above the actual demand curve in period 2, $p_2(x_2|\overline{x_1}, \widetilde{I}_1(\overline{x_1}))$. To avoid overcrowding the graph, we assume that it corresponds to the demand curve in period 2 when assuming away hysteresis, $p_{2,NoH}(x_2)$. The average marginal bias in period 2 is then represented by the trapezoid $B_2G_2D_2H_2$ in Figure B.5d. In any case, the total private welfare in period 2 is necessarily smaller (by the triangle $B_2D_2H_2$). Consequently, the welfare effect of the short-run policy decreases and the sign of the bias from assuming away hysteresis becomes ambiguous, depending on the relative sizes of the welfare gain from the correction of the externality and the loss in private welfare in

[80] Note that in a model with three periods, the demand curve in period 2 would further underestimates the gains in the household private welfare from reducing its electricity use in period 2.

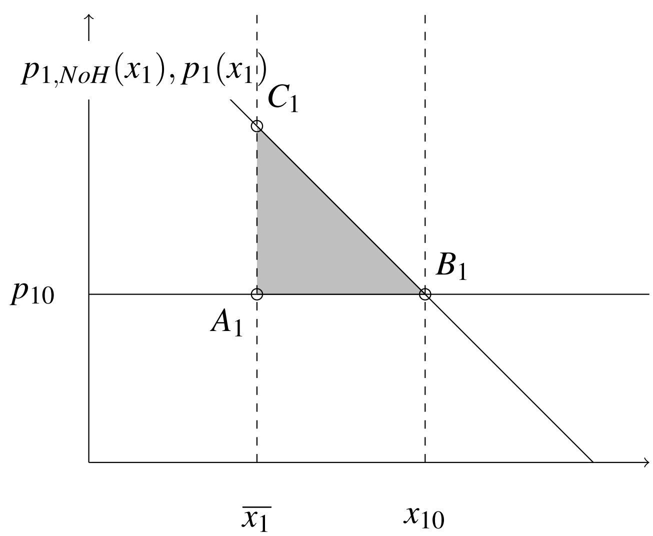
Figure B.5: The loss in private welfare from corrective policies with and without hysteresis in a model with myopic beliefs



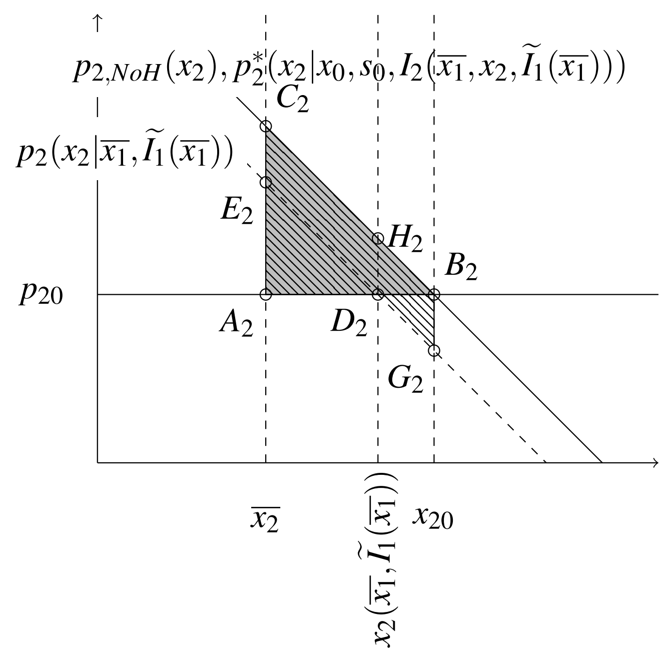
(a) Period 1 ("experienced" utility criteria)



(b) Period 2 ("experienced" utility criteria)



(c) Period 1 ("decision" utility criteria)



(d) Period 2 ("decision" utility criteria)

The figure illustrates the loss in private welfare in a model with a household who is myopic about the effect of its current choices on its future propensity to consume (the private welfare corresponds to the consumer surplus) for the short-run and the long-run corrective polices. Panels (a) and (b) assumes that the government evaluates private welfare based on the household's "experienced" utility. Panels (c) and (d) assumes that the government evaluates private welfare based on the household's "decision" utility. For simplicity, we assume linear demand curves and no change in the slope of the demand curve following a short-run policy in the presence of hysteresis. We also assume that the household expects to consume in period 2 what it would have consumed in absence of any corrective policy, $\widetilde{x_2} = x_{20}$. See text for details.

period 2. The long-run effect of the short-run policy remains key to estimate both of these effects.

Long-run policy. Let's consider now the long-run policy ($\overline{x_1} < x_{10}$, $\overline{x_2} < x_{20}$). Assuming away hysteresis, decisions in the two periods are assumed to be independent, so myopia is again inconsequential, and we obtain the same formula as in the benchmark framework with both criteria:

$$\begin{aligned}
\Delta W^{e,LR}_{NoH} &= \int_{x_{10}}^{\overline{x_1}} \frac{d\widetilde{V}(x_1,x_2,I_1,I_2)}{dx_1}dx_1 + \int_{x_{10}}^{\overline{x_1}} \left[\frac{dV(x_1,\widetilde{x}_2,\widetilde{I}_1(x_1),\widetilde{I}_2)}{dx_1} - \frac{d\widetilde{V}(x_1,x_2,I_1,I_2)}{dx_1}\right] dx_1 \\
&+\beta \int_{x_{20}}^{\overline{x_2}} \frac{dU_2(x_2,I_2)}{dx_2}dx_2 - \int_{x_{10}}^{\overline{x_1}} \frac{dE_1(x_1)}{dx_1}dx_1 - \beta \int_{x_{20}}^{\overline{x_2}} \frac{dE_2(x_2)}{dx_2}dx_2 \\
&= \int_{x_{10}}^{\overline{x_1}} \frac{dU_1(x_1,I_1)}{dx_1}dx_1 + \beta \int_{x_{20}}^{\overline{x_2}} \frac{dU_2(x_2,I_2)}{dx_2}dx_2 - \int_{x_{10}}^{\overline{x_1}} \frac{dE_1(x_1)}{dx_1}dx_1 - \beta \int_{x_{20}}^{\overline{x_2}} \frac{dE_2(x_2)}{dx_2}dx_2 \\
&= \int_{x_{10}}^{\overline{x_1}} [p_{1,NoH}(x_1) - p_{10}]\, dx_1 - \int_{x_{10}}^{\overline{x_1}} MD_1(x_1)dx_1 + \beta \int_{x_{20}}^{\overline{x_2}} [p_{2,NoH}(x_2) - p_{20}]\, dx_2 - \beta \int_{x_{20}}^{\overline{x_2}} MD_2(x_2)dx_2
\end{aligned} \tag{71}$$

$$\begin{aligned}
\Delta W^{d,LR}_{NoH} &= \int_{x_{10}}^{\overline{x_1}} \frac{d\widetilde{V}(x_1,x_2,I_1,I_2)}{dx_1}dx_1 + \beta \int_{x_{20}}^{\overline{x_2}} \frac{dU_2(x_2,I_2)}{dx_2}dx_2 - \int_{x_{10}}^{\overline{x_1}} \frac{dE_1(x_1)}{dx_1}dx_1 - \beta \int_{x_{20}}^{\overline{x_2}} \frac{dE_2(x_2)}{dx_2}dx_2 \\
&= \int_{x_{10}}^{\overline{x_1}} [p_{1,NoH}(x_1) - p_{10}]\, dx_1 - \int_{x_{10}}^{\overline{x_1}} MD_1(x_1)dx_1 + \beta \int_{x_{20}}^{\overline{x_2}} [p_{2,NoH}(x_2) - p_{20}]\, dx_2 - \beta \int_{x_{20}}^{\overline{x_2}} MD_2(x_2)dx_2
\end{aligned} \tag{72}$$

It is useful to consider similar steps as before to derive a formula for the welfare effect when the researcher allows for hysteresis. The first step captures the effect of the policy, as if there was only a short-run policy. For this step, let's define again $\widetilde{x}_2$, $\widetilde{I}_1(x_1)$, and $\widetilde{I}_2$ as the household (expected) choices according to equations (63) and (64) for a given value of x_1. Let's also still define $x_2(\overline{x_1},\widetilde{I}_1(\overline{x_1}))$ and $I_2(\overline{x_1},x_2,\widetilde{I}_1(\overline{x_1}))$ as the actual choices in period 2 given the first-order conditions (66). The second step captures the effect of the period-2 policy, holding fixed investment levels from period 1. For this step, let's define $\widetilde{I}_2(\overline{x_1},x_2,\widetilde{I}_1(\overline{x_1}))$ as the investment choices in period 2 according to equation (66) for a given value of x_2 and of the choice variables in period 1. Finally, the third step in the benchmark framework is absent in this model because the myopic household will not take any action in anticipation of the period-2 policy, such as choosing different investment levels. The welfare effect of the policy is then:

$$
\begin{aligned}
\Delta W^{e,LR} =& \int_{x_{10}}^{\overline{x_1}} \frac{d\widetilde{V}(x_1,x_2,I_1,I_2)}{dx_1}dx_1 + \int_{x_{10}}^{\overline{x_1}} \left[\frac{dV(x_1,\widetilde{x_2},\widetilde{I_1}(x_1),\widetilde{I_2})}{dx_1} - \frac{d\widetilde{V}(x_1,x_2,I_1,I_2)}{dx_1}\right] dx_1 \\
&+\beta \int_{\widetilde{x_2}}^{x_2(\overline{x_1},\widetilde{I_1}(\overline{x_1}))} \frac{dU_2(\overline{x_1},x_2,\widetilde{I_1}(\overline{x_1}),I_2)}{dx_2}dx_2 + \beta \int_{x_2(\overline{x_1},\widetilde{I_1}(\overline{x_1}))}^{\overline{x_2}} \frac{dU_2(\overline{x_1},x_2,\widetilde{I_1}(\overline{x_1}),I_2)}{dx_2}dx_2 \\
&- \int_{x_{10}}^{\overline{x_1}} \frac{dE_1(x_1)}{dx_1}dx_1 - \beta \int_{x_{20}}^{\overline{x_2}} \frac{dE_2(x_2)}{dx_2}dx_2 \\
=& \int_{x_{10}}^{\overline{x_1}} [p_1(x_1) - p_{10}]\, dx_1 + \int_{x_{10}}^{\overline{x_1}} \left[p_1^*(x_1|\widetilde{x_2},\widetilde{I_1}(x_1),\widetilde{I_2}) - p_1(x_1)\right] dx_1 \\
&+\beta \int_{\widetilde{x_2}}^{x_2(\overline{x_1},\widetilde{I_1}(\overline{x_1}))} \left[p_2(x_2|\overline{x_1},\widetilde{I_1}(\overline{x_1})) - p_{20}\right] dx_2 + \beta \int_{x_2(\overline{x_1},\widetilde{I_1}(\overline{x_1}))}^{\overline{x_2}} \left[p_2(x_2|\overline{x_1},\widetilde{I_1}(\overline{x_1})) - p_{20}\right] dx_2 \\
&- \int_{x_{10}}^{\overline{x_1}} MD_1(x_1)dx_1 - \beta \int_{x_{20}}^{\overline{x_2}} MD_2(x_2)dx_2 \qquad (73)
\end{aligned}
$$

$$
\begin{aligned}
\Delta W^{d,LR} =& \int_{x_{10}}^{\overline{x_1}} \frac{d\widetilde{V}(x_1,x_2,I_1,I_2)}{dx_1}dx_1 + \beta \int_{\widetilde{x_2}}^{x_2(\overline{x_1},\widetilde{I_1}(\overline{x_1}))} \frac{dU_2(\overline{x_1},x_2,\widetilde{I_1}(\overline{x_1}),I_2)}{dx_2}dx_2 + \beta \int_{x_2(\overline{x_1},\widetilde{I_1}(\overline{x_1}))}^{\overline{x_2}} \frac{dU_2(\overline{x_1},x_2,\widetilde{I_1}(\overline{x_1}),I_2)}{dx_2}dx_2 \\
&+\beta \int_{\widetilde{x_2}}^{\overline{x_2}} \left[\frac{dU_2(s_0,x_2,\widetilde{I_2}(\overline{x_1},x_2,\widetilde{I_1}(\overline{x_1})))}{dx_2} - \frac{dU_2(\overline{x_1},x_2,\widetilde{I_1}(\overline{x_1}),I_2)}{dx_2}\right] dx_2 \\
&- \int_{x_{10}}^{\overline{x_1}} \frac{dE_1(x_1)}{dx_1}dx_1 - \beta \int_{x_{20}}^{\overline{x_2}} \frac{dE_2(x_2)}{dx_2}dx_2 \\
=& \int_{x_{10}}^{\overline{x_1}} [p_1(x_1) - p_{10}]\, dx_1 + \beta \int_{\widetilde{x_2}}^{x_2(\overline{x_1},\widetilde{I_1}(\overline{x_1}))} \left[p_2(x_2|\overline{x_1},\widetilde{I_1}(\overline{x_1})) - p_{20}\right] dx_2 \\
&+\beta \int_{x_2(\overline{x_1},\widetilde{I_1}(\overline{x_1}))}^{\overline{x_2}} \left[p_2(x_2|\overline{x_1},\widetilde{I_1}(\overline{x_1})) - p_{20}\right] dx_2 + \beta \int_{\widetilde{x_2}}^{\overline{x_2}} \left[p_2^*(x_2|s_0,I_2(\overline{x_1},x_2,\widetilde{I_1}(\overline{x_1}))) - p_2(x_2|\overline{x_1},\widetilde{I_1}(\overline{x_1}))\right] dx_2 \\
&- \int_{x_{10}}^{\overline{x_1}} MD_1(x_1)dx_1 - \beta \int_{x_{20}}^{\overline{x_2}} MD_2(x_2)dx_2 \qquad (74)
\end{aligned}
$$

There are two main differences with the benchmark framework when evaluating private welfare based on the household's experienced utility, which are the same as for the short-run policy. First, the second term in equation (73) captures the average marginal bias when changing consumption in period 1 from x_{10} to $\overline{x_1}$, holding fixed the expected choices of the other variables according to equations (63) and (64). This empirical object, which corresponds to the trapezoid $B_1C_1F_1G_1$ in Figure B.5a, is not easily estimated but it is again necessarily positive because the policy corrects the internality issue in the right direction. Second, the third term in equation (73) captures the increase in private welfare from the period-1 policy when the household revises its consumption level in period 2 compared to expectations, from $\widetilde{x_2}$ to $x_2(\overline{x_1},\widetilde{I_1}(\overline{x_1}))$. This corresponds to the triangle $B_2D_2G_2$ in Figure B.5b. The other terms in equation (73) are similar as in the benchmark framework. The bias from assuming away hysteresis (abstracting from differences in the slopes of the demand curves) thus corresponds to the trapezoids $B_1C_1F_1G_1$ and $B_2C_2E_2G_2$ in Figures B.5a and B.5b, respectively. The long-run effect of the short-run policy, $x_{20} - x_2(\overline{x_1}, \widetilde{I}_1(\overline{x_1}))$, remains a key empirical object to estimate in order to evaluate the bias in period 2 (the trapezoid $B_2C_2E_2G_2$).

There are two main differences with the benchmark framework when evaluating private welfare based on the household's decision utility in period 1. The first difference comes from the second term in equation (74), which is the same as the second term in equation (70). It corresponds to the change in consumption levels in period 2 caused by the period-1 policy, which the household perceives as increasing its private welfare, or the triangle $B_2D_2G_2$ in Figure B.5d. The third term in equation (74) is the same as in the benchmark framework and corresponds to the perceived loss in private welfare from the additional reduction in electricity use due to the period-2 policy. As before, it corresponds to the triangle $A_2D_2E_2$. The second difference then comes from the fact that the perceived change in marginal utility from reducing electricity use below expected level $\widetilde{x_2}$ is inaccurate according to this welfare criterion. The fourth term in equation (74) captures the associated average marginal bias, the difference between the accurate valuation and the demand curve, between expected and actual consumption levels in period 2. This is captured by the trapezoid $B_2C_2E_3G_2$ in Figure B.5d, in which we assume again $\widetilde{x_2} = x_{20}$ and $p_2^* = p_{2,NoH}(x_2)$ for simplicity. In the specific case displayed in Figure B.5d, there is at the end no bias from assuming away hysteresis because the resulting overall loss in private welfare corresponds to triangle $A_2B_2C_2$. Of course, the accurate valuation does not have to correspond to the one assumed when neglecting the possibility of hysteresis as in Figure B.5d. The point is made, however, that there may be no bias from assuming away hysteresis when evaluating private welfare based on the household's decision utility in period 1. This is intuitive: if the shift of the demand curve in period 2 only represents a perceived change in the valuation of x_2 and not an actual change, there is no reason for the long-run effect of a short-run policy to imply a smaller loss in private welfare from a corrective policy in period 2.

In sum, the bias from assuming away hysteresis is positive and may even be larger in this model using the household's "experienced" utility, which would make sense in many applications. However, using the decision utility in some prior period could make sense with, e.g., addictions.

C The causes of the 2001 Brazilian electricity crisis: additional information

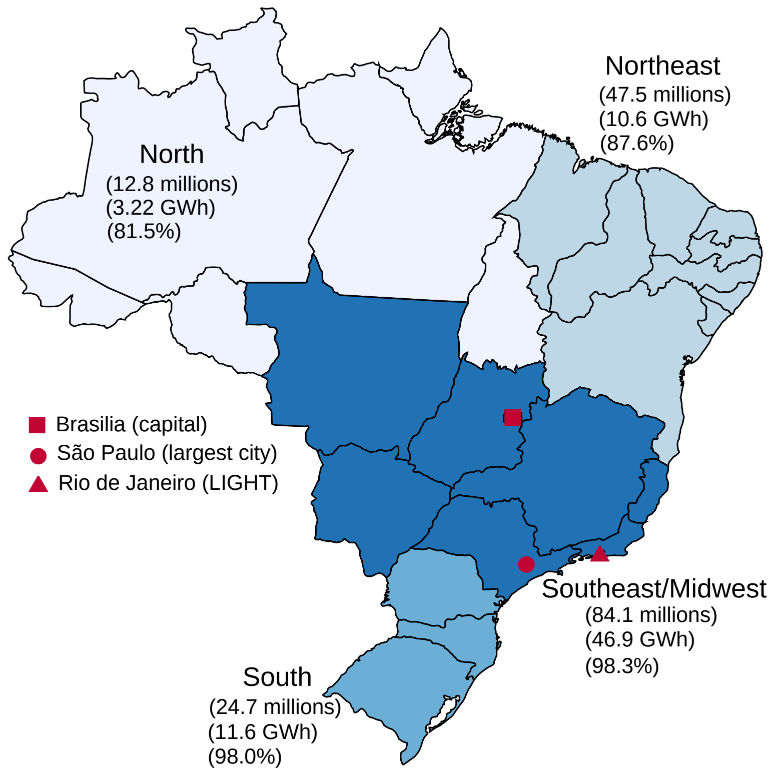
This section presents additional information on the causes of the 2001 Brazilian electricity crisis.

Figure C.1 presents the map of Brazil, highlighting the four subsystems of the National Interconnected System of Brazil with the population, total residential electricity demand and the share of households connected to electricity, all values of 2000.

Figure C.2 displays the evolution of hydro–reservoirs' capacity in percentage of their maximum capacity over calendar months within each year between 1991 and 2011 in the Southeast/Midwest (panel a) and in the South (panel b). These are in the same format as Figures 3b and 3c in the text of the paper, which display the streamflow level of the rivers serving the reservoirs in the two regions. The solid line corresponds to 2001 (the year the electricity crisis started), the dashed line to 2000, and the dotted lines to all other years. Figure C.2a shows the clear seasonal pattern in the Southeast/Midwest, with heavy rainfall upstream of the rivers that serve the reservoirs replenishing them at the beginning of every year. Figures C.2a and C.2b also show that the levels of the reservoirs were very low in both regions at the beginning of 2000. Next, Figures C.2a shows that the streamflow patterns displayed in Figures 3b – that streamflow levels were much lower than average at the beginning of 2001 in the Southeast/Midwest – implied that the level of the reservoirs did not increase in the Southeast/Midwest in 2001, as they usually do at the beginning of every year. In contrast, Figures C.2b shows that the streamflow patterns displayed in Figures 3c – that streamflow levels were higher than average at the end of 2000 in the South – implied that the reservoirs were completely replenished at the start of 2001 in the South.

Table C.1 presents the realized electricity demand in each subsystem and year as a percentage of the demand forecast from the 1997-2007 Decennial Energy Plan (PDE) produced by the National System Operator along with the Mining and Energy Ministry. This is the main national plan that guides the medium- and long-run expansion of energy infrastructure in the country. We can see in the first cell in column (1) that the energy used in the Southeast in 1998 was 99.6 percent of the forecast demand (in PDE 1997) for that region and year. We can see in the table that the actual growth in demand never outpaced growth in projected demand. However, it systematically outpaced the growth in generation capacity in the years prior to 2001 and the crisis could have been avoided if generation capacity had been expanded adequately according to the plans – e.g., if several infrastructure projects had not been delayed or canceled (Kelman, 2001). See Kelman (2001), Maurer, Pereira and Rosenblatt (2005), and Mation and Ferraz (2011) for more discussion on the cause of the crisis and the exogenous role of weather in the differential treatment across subsystems.

Figure C.1: Map of the 4 Subsystems of the National Interconnected System of Brazil



This map presents the four subsystems of the National Interconnected System of Brazil. The first number in parentheses is the population, the second number is the total residential energy consumption in 2000, and the third is the percentage of households connected to electricity. The three red markers locate the two largest cities in Brazil (São Paulo and Rio de Janeiro) and the capital (Brasilia). Source: Census 2000 and National System Operator (ONS).

Figure C.2: Stocked energy of the reservoirs by calendar month

Official data from ONS (the National System Operator). The figure displays the evolution of hydro–reservoirs' capacity in percentage of their maximum capacity over calendar months within each year between 1991 and 2011 in the Southeast/Midwest (panel a) and in the South (panel b). The graphs are in the same format as those in Figures 3b and 3c in the text of the paper, which display the streamflow level of the rivers serving the reservoirs in the two regions. The solid line corresponds to 2001 (the year the electricity crisis started), the dashed line to 2000, and the dotted lines to all other years. Together with Figure 3, this figure shows that the crisis and the differential treatment between subsystems were due to a weather shock (and not to a demand shock).

Table C.1: Realized Electricity Demand as Percentage of Demand Forecast (%)

	Southeast	**Midwest**	**South**	**Brazil**
	(1)	(2)	(3)	(4)
1998	99.6	98.5	97.9	99.4
1999	95.6	96.4	97.5	95.6
2000	96.2	95.7	98.5	95.6

Source: Forecasts from 1997-2007 Decennial Energy Plan (PDE) produced by the National System Operator (ONS) along with the Mining and Energy Ministry. Realized demand from ONS.

D Timeline of the electricity crisis: additional information

This section provides more details on the timeline of the electricity crisis.

Late 1999 — The National System Operator (ONS) presents simulations of hydrological scenarios for 2000 based on the actual reservoir levels on November 30, 1999. The report concludes that the reservoir levels in some regions would hit zero (i.e., no electricity) in 13% of these scenarios. (*ONS-DPP 059/1999*)

Feb 2000 — The Ministry of Mining and Energy (MME) creates the Priority Thermal Program (PPT) to increase the generation capacity of thermal power plants as the "unique solution" to a possible collapse of the system.

Early 2000 — The Priority Thermal Program becomes the Emergency Thermal Program.

Jul 2000 — In a meeting with the president and his economic advisors, the minister of the MME dismisses the chances of any energy crisis during 2000-2003.[81]

Dec 2000 — ONS forecasts a scenario for 2001 with no energy crisis.

Feb 2001 — Hydrological conditions reach 70% of the long run average, and ONS radically changes the forecast for 2001.

Mar 2001 — ONS officially requests that the federal government intervene to assure a 20% load reduction.

Mar 2001 — *For the first time, the regulatory agency (ANEEL) publicly addresses a possible imminent electricity shortage*. It proposes a Consumption Reduction and Supply Increase Plan (RECAO), which was abandoned shortly afterward.

Apr 2001 — The Priority Thermal Program (PPT) fails and MME starts designing an incentive-based load reduction program.[82]

[81] Based on documents from the National System Operator (ONS), the minister stated: "considering the Priority Thermal Program (PPT), even if we observe an increase in demand larger than expected, we will not face energy supply and peak problems during 2000-2003 as long as the hydrological conditions are above 85% of the long run average".

[82] (*O Globo*, April 23, 2001). "Plan to hold expenditure on electricity" aims to reduce consumption in three regions with 25 measures. In case these measures are not effective, it is possible that these regions will have blackouts in June. (*Folha de São Paulo*, Front page, A1, 06/04/2001). "Plan to avoid energy saving program failed." Only three of the planned measures were implemented. (*Folha de São Paulo*, B7, 05/05/2001)

May 2001 The government announces a temporary energy-saving program to be implemented on June 4. This announcement receives a lot of attention from the media.[83]

Jun 2001 The temporary energy-saving program is implemented and, from the very start, is expected to last until February 2002, the end of the next rainy season (*Veja*, July 19, 2001).

Feb 2002 Household fines and threats of electricity cuts are suspended.[84]

March 2002 Last billing cycle (February-March) in which bonuses were paid.

[83] *Folha de São Paulo*: "Government has not decided between regular supply interruptions or higher tariffs" (Front page, A1, 15/05/2001); "Plan will affect households with electricity bills above R$29" (Front page, A1, 18/05/2001); "Government imposes 'super tariffs' and will cut electricity of those who don't save" (Front page, A1, 19/05/2001); "Households should avoid storing food at home and shop for groceries more often" (B10, 29/05/2001); "Subsidies do not reduce lightbulbs' prices" (B7, 01/06/2001).

[84] "Rain brings relief to reservoirs" (*Folha de São Paulo*, B1, 03/01/2002).

E The energy-saving program: additional information

This section presents additional information related to the energy-saving program of the 2001 Brazilian electricity crisis.

Figure E.1 provides an example of the letter received by residential customers at the start of the energy-saving program informing them about their quota.

Figure E.2 explains the rules that the energy-saving program used to assign individual quotas to customers at the beginning of the crisis. Customers' *baseline* was defined as the average billed monthly consumption from May to July 2000 (or the first three monthly bills for customers who moved in after May 2000). Quotas were set at 80% of the baseline, with three exceptions: (i) customers with a baseline below 100 kWh had their quotas set at 100% of baseline; (ii) customers with a baseline above 100 kWh whose quotas would have been below 100 kWh using the 80% rule had their quotas set at 100 kWh; (iii) because quotas were based on billed consumption and bills always charge for minimum consumption levels, quotas were at least equal to these minimum levels. Figure E.2 illustrates the case of LIGHT, the distribution utility serving the city of Rio de Janeiro and surrounding municipalities, where minimum levels are 30 kWh, 50 kWh, and 100 kWh for monophasic, biphasic, and triphasic connections, respectively.

Figure E.3 provides an example of how the pecuniary incentives of the energy-saving program modified the cost of consuming electricity during the crisis. The figure considers the case of customers with a quota of 250 kWh (80% of baseline in the first five months of the crisis, before any change in quotas). We assume a budget of R$500 (in 2001 $\simeq$ US\$576.9 in 2012) and a tariff of R$.208/kWh (main tariff for LIGHT residential customers in June 2001 $\simeq$ US\$.24 in 2012). The usual marginal cost of electricity is nil up to 100 kWh because we assume a minimum consumption level of 100 kWh (triphasic connection). During the crisis, the total cost of electricity is nil if consuming below 100 kWh because of the guaranteed bonus. Conditional on exceeding the quota, the cost of electricity increases because of the fines paid for every kWh above 200. Above the quota, the fines (i) increase the marginal price (by 50% up to 500 kWh, then by 200%) and (ii) increase the cost discretely by $(250-200) \times .208 \times 50\%$ =R$5.2 ($\simeq$ US$6 in 2012).

Figure E.1: Example of a letter informing residential customers about their quota at the start of the energy-saving program

Light

Rio de Janeiro, 30 de maio de 2001

Prezado Cliente,

Atendendo à Resolução nº 4 da Câmara de Gestão da Crise de Energia Elétrica, a Light informa a sua meta de consumo:

137 kWh/mês

De acordo com a mesma Resolução, a partir do dia 04 de junho, os consumidores que ultrapassarem suas metas ficarão sujeitos à suspensão do fornecimento de energia elétrica.

Assim, fique atento ao seu consumo e lembre-se que, se ele for menor do que a sua meta, você poderá ter direito a um bônus.

Com a sua participação e a participação de todo o Rio de Janeiro, vamos enfrentar melhor o desafio do racionamento. Isso tem um nome:

Energia Solidária.

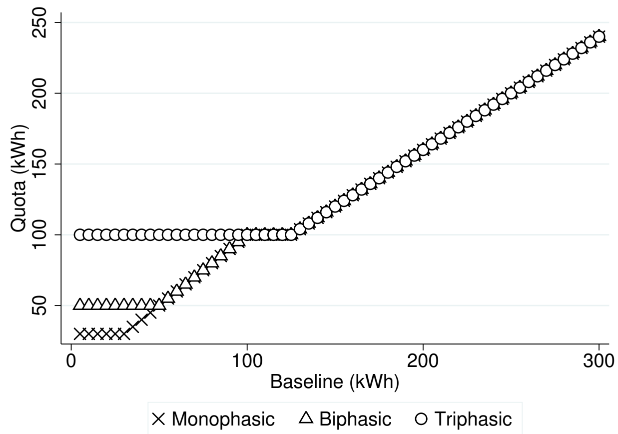
Para maiores informações sobre recionamento:

- 0800-282-0120
- Agências Light
- www.lightrio.com.br

Desde já, agradecemos a sua compreensão.

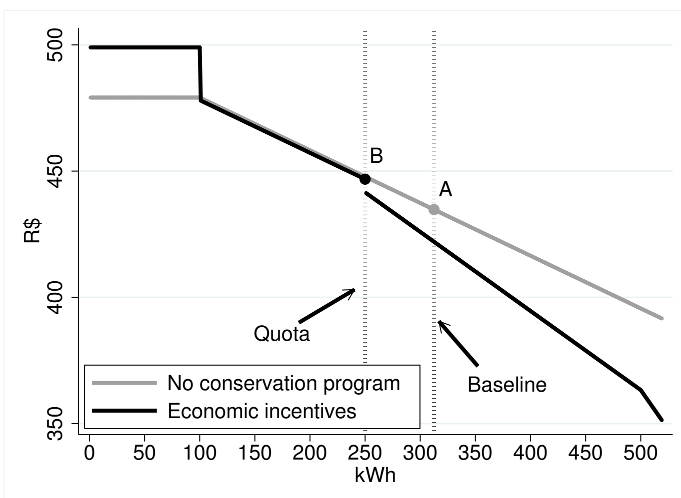
Light Serviços de Eletricidade S.A.

Figure E.2: Quota assignment rule for the energy-saving program



The figure explains the rule used in the energy-saving program to assign individual quotas to customers at the beginning of the crisis. Customers' *baseline* was defined as the average billed monthly consumption from May to July 2000 (or the first three monthly bills for customers who moved in after May 2000). Quotas were set at 80% of the baseline, with three exceptions: (i) customers with a baseline below 100 kWh had their quotas set at 100% of baseline; (ii) customers with a baseline above 100 kWh whose quotas would have been below 100 kWh using the 80% rule had their quotas set at 100 kWh; (iii) because quotas were based on billed consumption and bills always charge for minimum consumption levels, quotas were at least equal to these minimum levels. The figure illustrates the case of LIGHT, the distribution utility serving the city of Rio de Janeiro and surrounding municipalities, where minimum levels are 30 kWh, 50 kWh, and 100 kWh for monophasic, biphasic, and triphasic connections, respectively.

Figure E.3: Example of the economic incentives of the energy-saving program



The figure displays the pecuniary incentives of the energy-saving program for customers with a quota of 250 kWh (80% of baseline). The monetary values are in R\$2000 (with R\$1 in 2000 $\simeq$ 1.154 US\$ in 2012). We assume a budget of R\$500 and a tariff of R\$.208/kWh (main tariff for LIGHT residential customers in June 2001. The usual marginal cost of electricity is nil up to 100 kWh because we assume a minimum consumption level of 100 kWh (triphasic connection). During the crisis, the total cost of electricity is nil if consuming below 100 kWh because of the guaranteed bonus. Conditional on exceeding the quota, the cost of electricity increases because of the fines paid for every kWh above 200. Above the quota, the fines (i) increase the marginal price (by 50% up to 500 kWh, then by 200%) and (ii) increase the cost discretely by $(250-200)\times .208\times 50\%$ =R\$5.2.

F Data description: additional information

This section presents additional information on the various datasets used in the paper and on the process of preparing these datasets for analysis.

A. ANEEL administrative data – Electricity use.

The monthly data on electricity use at the level of distribution utilities come from two datasets provided by ANEEL. The first one spans a period from January 1991 to April 2005; the second one spans a period from January 2003 to August 2015. The data include seven variables for every distribution utility in every month: name of the distribution utility, year, month, customer category, total number of customers, total consumption (kWh billed), and total revenue. There are eight customer categories: residential, commercial, industrial, rural, own consumption, public lighting, public services, and government.

We used these data to construct two main outcomes of interest: the average monthly electricity use per customer (total consumption/total number of customers) and the average monthly electricity price (total revenue/total consumption) by customer category (we focus on residential customers). The data are publicly available but must be requested from ANEEL.

We used the overlapping period (January 2003 to April 2005) to check that the two series are fully consistent. A very few inconsistent entries were corrected by hand (e.g., if total consumption doubles from one month to the next, we replace the value of the observation by the average of the observation in the previous and subsequent month; our codes are available upon request).

The concession areas of a few distribution utilities were divided over time, giving birth to new distribution utilities. Therefore, when we use data on distribution utilities starting in 2000 (resp. in 1996 or in 1991) in the analysis, we aggregate the data according to the concession areas as defined in 2000 (resp. in 1996 or in 1991).

B. ANEEL administrative data – Electricity tariffs.

The legal document for every price-setting regulation (regulations are periodic – usually yearly – according to the terms of concession of each distribution utility) can be found in the online version of the Diário Oficial da União (and on ANEEL's website starting in 2004). We gathered copies of every regulation since 1996 (when the electricity sector was privatized) and entered the data on the electricity tariffs for every customer category in spreadsheets, including the first and last month that these tariffs were in application. From 2005 onward, the regulations also specify the "exogenous" cost-of-energy component for every tariff, which distribution utilities are allowed to pass onto the tariffs.

C. Other utility-level variables.

We matched the above data at the level of distribution utilities to decennial census data (2000 and 2010), and yearly data on population (1992-2014), GDP (1999-2012), formal employment, and average temperature (1996-2014). ANEEL provided us with a list of all the municipalities within the concession area of every distribution utility; we use that information to aggregate the information in these datasets to the level of each distribution utility.

We obtained the microdata of the decennial censuses conducted in 2000 and 2010 from the Brazilian Institute of Geography and Statistics (IBGE). We obtained population estimates by municipality and year created by IBGE on IPEADATA, which is the data platform of IPEA, an important research institution of the Brazilian government. The population estimates were available on IPEADATA from 1992 to 2014. They were missing in 3 years; we extrapolated the missing data by taking the average of the observation in the previous and subsequent year for the same municipality. We obtained GDP estimates by municipality and year created by IBGE on IBGE's website, where they were available from 1999 to 2012. Data on the number of formal workers by municipality and year comes from the RAIS dataset. RAIS is a longitudinal matched employee-employer dataset covering by law the universe of formally employed workers, including public employees. RAIS is not publicly available, but it can be requested for research purposes from the Brazilian Ministry of Labor. We obtained a version of RAIS covering the period from 1996 to 2014. RAIS contain information on the start and end of every formal employment spell in every year (an employee-employer match). We used this information to create a variable capturing the total number of formal workers by municipality and month in every year from 1996 to 2014. The data on average temperature come from Matsuura and Willmott (2012). These are monthly-mean air temperatures at the surface (land-only) on a grid of resolution 0.5 by 0.5 (centered on 0.25 degree). We take the average monthly-mean air temperature by municipality and year. The monthly-mean temperature by distribution utility is then the weighted average of the monthly-mean air temperature of the municipalities in the concession area of each distribution utility, where we use the municipalities' population as weights.

We also use the consumer price index created by IBGE (IPCA) to express nominal values in real terms. The consumer price index contains monthly indices for the main metropolitan areas in each region. Unfortunately, Brazil does not have a consumer price index for the whole country.

The data on the streamflow in the rivers that serve hydropower plants (1996-2015) and on the hydro-reservoirs' water level (1991-2015) were obtained on the website of the National System Operator (ONS), the body responsible for running the electricity generation and transmission systems in Brazil. The data are aggregated by subsystem and are available at the monthly level.

D. Household-level billing data (LIGHT).

We obtained household-level billing data for every low-voltage customer from one distribution utility, LIGHT. The data are not publicly available and were obtained for this specific research project. The data that were provided to us consist of five monthly registries from January 2000 to December 2005.

The *client* registry includes 10 fields: period identifier (year and month), client identifier, address identifier (2 fields), including zip code, and five fields (strings) for the location of the housing unit.

The *metering* registry includes 12 fields: period identifier (year and month), invoice identifier, new reading of the meter, new reading date, previous reading, previous reading date, a coefficient to convert readings into kWh (depending on meter type), difference between the two readings in kWh, number of days between the two readings, average consumption between the two readings, connection type (monophasic, biphasic, triphasic), reading type (interior, exterior, estimated).

The *invoice* registry includes 6 fields: period identifier (year and month), invoice identifier, client type (e.g., residential main tariff, residential alternative tariff, commercial, rural), invoice creation date, client identifier, invoice value.

The *detailed invoice* registry includes every invoice component: period identifier (year and month), invoice identifier, a code identifying the specific invoice element (e.g., fine or bonus during the crisis), a code identifying the tariff/price category associated with the specific invoice element, the quantity (e.g., metered kWh), the prevailing price/tariff, and the value associated with the specific invoice element.

The *crisis* registry includes 6 fields during the months of the electricity crisis: period identifier (year and month), invoice identifier, the prevailing quota adjusted for the number of days between the two readings, the quota originally assigned to each customer, and two fields capturing any trade in quotas between low-voltage industrial firms.

In June 2001, there were about 2,615,300 residential customers on the main tariff, and 482,800 residential customers on the alternative tariff. We restrict our sample to residential consumers (codes FBT100 and FBT180). We only keep bills with readings between 11 and 45 days apart. Since metering happens in different dates, we assign each metering to a given month if more than half of the metered days were in a given month. For example, consumption metered between March 20 and April 20 is assigned to April. We create a normalized measure of monthly consumption by dividing the metered consumption by the number of days between readings and multiplying by 30. We define neighborhood using the 5-digit zip code (Brazilian zip codes have 8 digits).

To construct the *balanced panel*, we selected residential customers that satisfied the following conditions: (i) paid regular tariffs (code FBT100) in the 3 months before the crisis (March to May 2001); (ii) first metering in Jan/2000 and last metering in Nov or Dec/2005; and (iii) have less than 10 (out of 72) missing metering month in the period.

To construct the *movers sample*, we selected residential customers that satisfied the following conditions: (i) paid regular tariffs (code FBT100) in the 3 months before the crisis (March to May 2001); (ii) first metering after Mar/2000 (inclusive) and last metering after Dec/2002 or Dec/2005 (depending on the sample); (iii) metered before the crisis (Marh-May 2001), during the crisis (July-

December 2001), and after the crisis (July-December 2002, 2003, 2004 and 2005); and (iv) have more than 6 non-missing metering until Dec/2002. We used the same procedure to construct the movers samples used in the placebo exercises, only changing the moving years as explained in the Appendix table.

E. Appliances and Habits of Use Surveys (PPH).

The PPH surveys are based on a representative sample of residential customers (legally connected to the grid) from several utilities, who were surveyed between July 1998 and June 1999 (first round) and between July 2004 and June 2005 (second round). The microdata are not publicly available but can be obtained for research purposes from PROCEL. The data were recorded during in-house interviews, which included questions on household characteristics, appliance ownership, and consumption habits. The sampling scheme did not change between rounds, but there was no attempt to interview the same households, so the data consist of cross-sections in each round. We use data from the ten distribution utilities – eight in the Southeast/Midwest and two in the South – that were surveyed in both rounds.[85] Conveniently, our estimated long-term impact on residential electricity use is stable starting around the time of the second PPH survey round (see Figure 4).

We input the average monthly electricity use of appliances in 1999 (before the crisis) 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 (see Table A.5). In the Appendix, appliances frequently use stand for appliances utilized more than four times a week. Shower thermostats have three power settings: Off, Low Power (*Mode Verão*), or High Power (*Modo Inverno*). A shower regulated in Low Power consumes 30% less energy than one regulated in High Power.

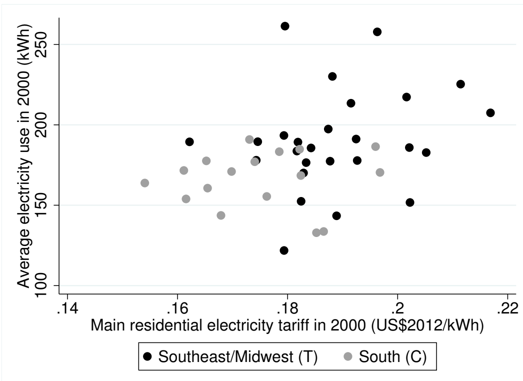
[85]We were not given the identity of those distribution utilities due to confidentiality concerns (but we were told that Light was one of them). Therefore, we cannot match the PPH survey data to the ANEEL administrative data.

G The joint distribution of average electricity use and covariates before the crisis (2000) for distribution utilities in the Southeast/Midwest and in the South

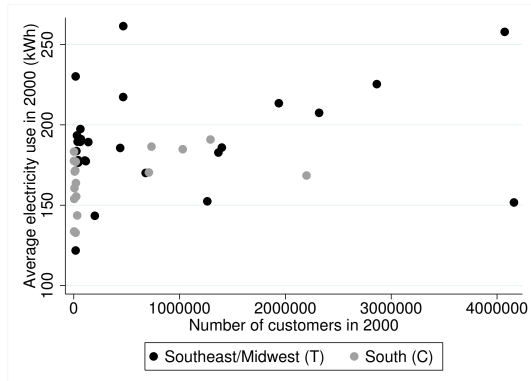
The following figures display the joint distribution of monthly average residential electricity use per customer and relevant covariates before the crisis for distribution utilities in the Southeast/Midwest and in the South. We present figures for all the variables displayed in Tables 1 and A.1. The data come from either the utility-level data from ANEEL or from the 2000 census matched to the concession area of each distribution utility. The figures show that there was some overlap in the distributions of average residential electricity use per customer in the Southeast/Midwest and in the South before the crisis. They also show that there was some overlap in the distributions of almost all covariates in the Southeast/Midwest and in the South before the crisis. The exception is for average temperatures, which are lower in the South and for which there is only limited overlap.

Figure G.1: Joint distribution of average electricity use and relevant covariates before the crisis I (2000)

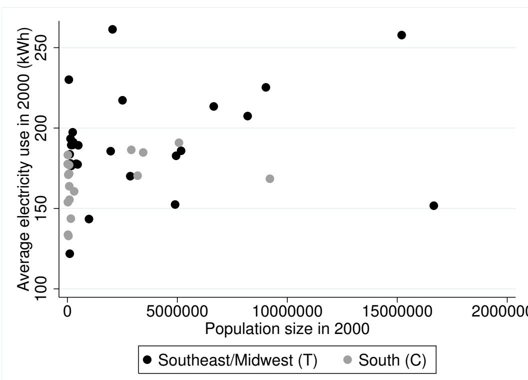
(a) Main residential electricity tariff



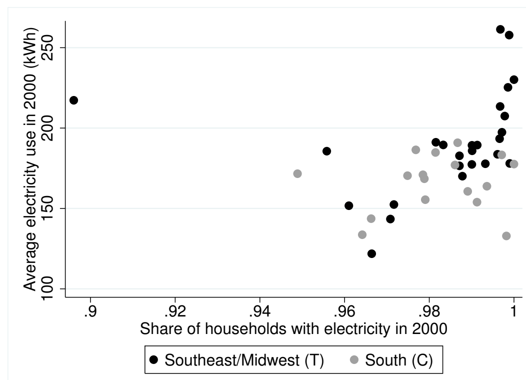
(b) Number of customers



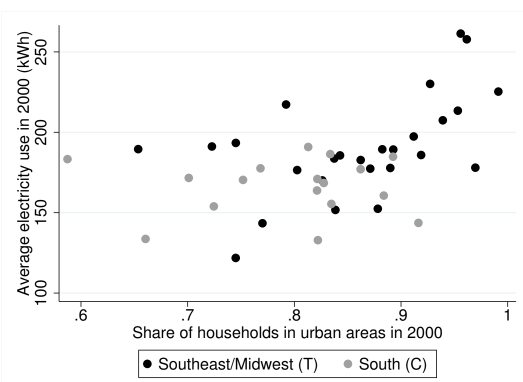
(c) Population size



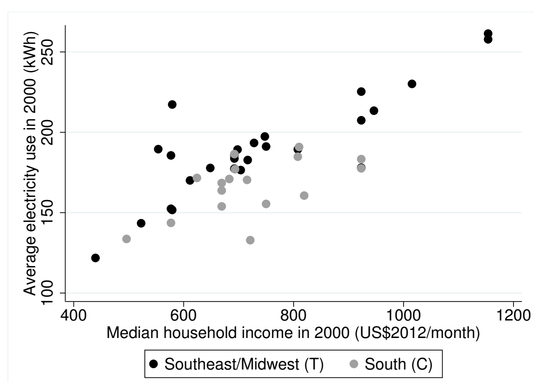
(d) Share with electricity



(e) Share urban

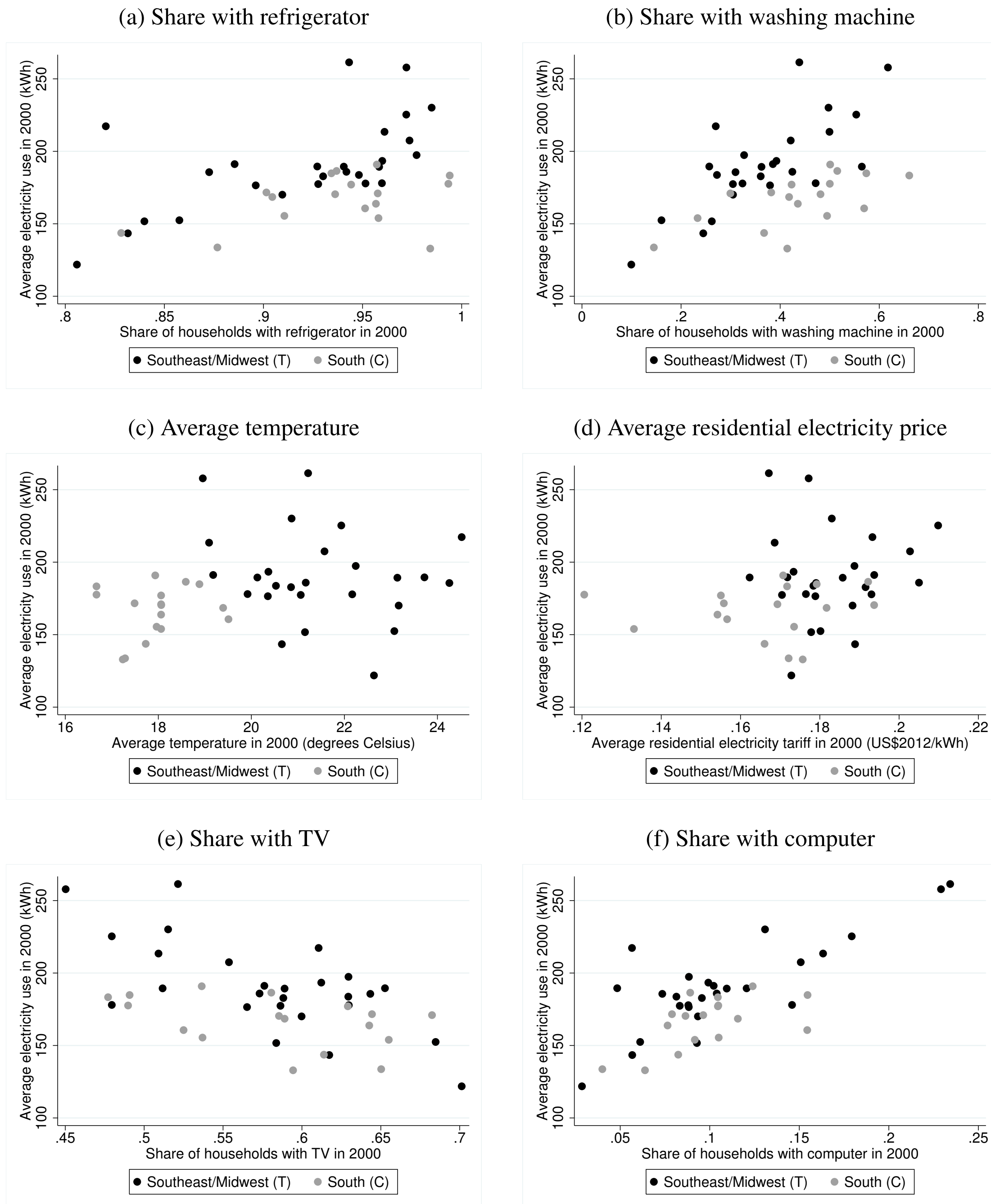


(f) Median household income



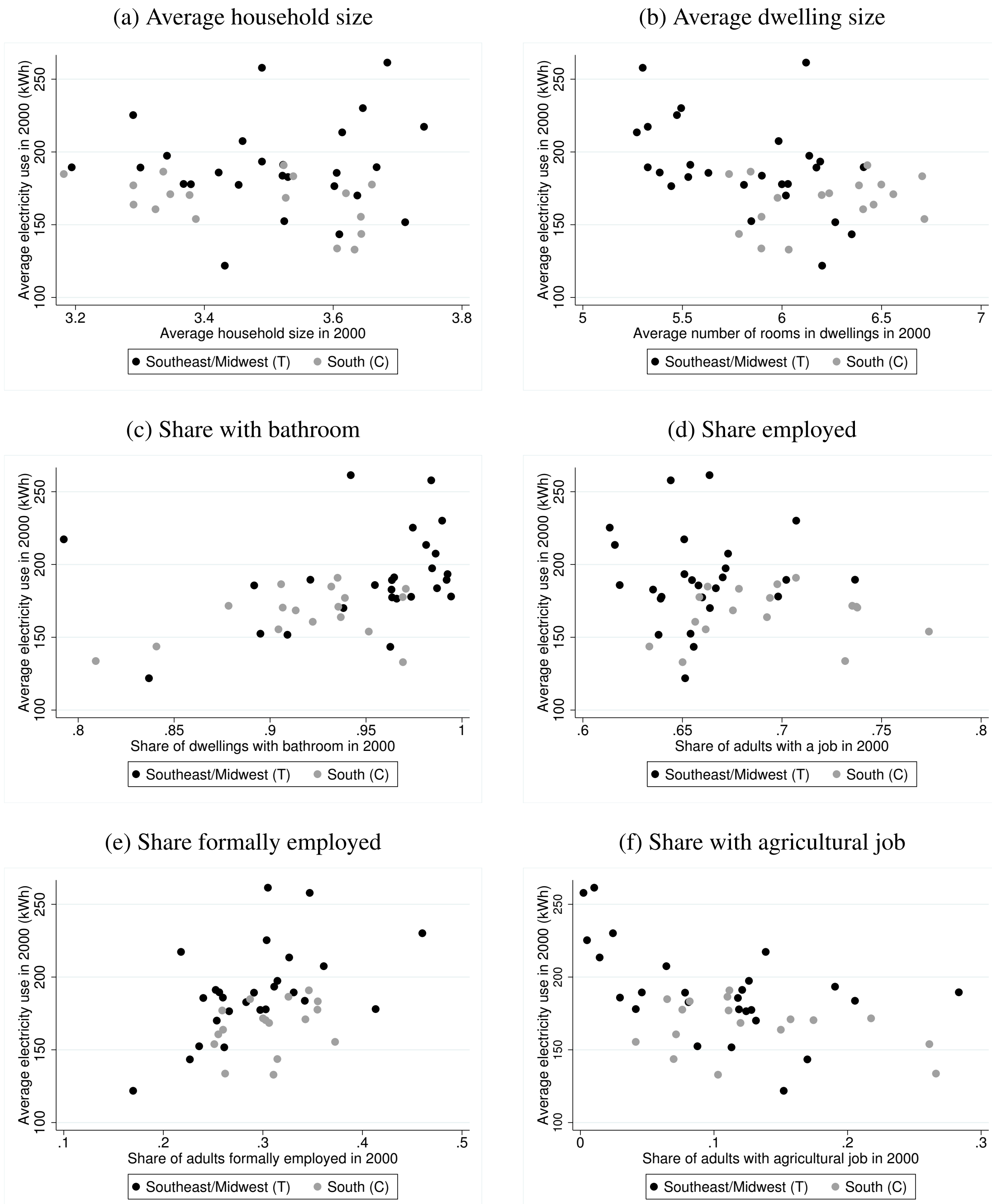
Each observation corresponds to a distribution utility and its concession area. The panels display the joint distribution of monthly average residential electricity use per customer and relevant covariates before the crisis (2000) for distribution utilities in the Southeast/Midwest and in the South.

Figure G.2: Joint distribution of average electricity use and relevant covariates before the crisis II (2000)



Each observation corresponds to a distribution utility and its concession area. The panels display the joint distribution of monthly average residential electricity use per customer and relevant covariates before the crisis (2000) for distribution utilities in the Southeast/Midwest and in the South.

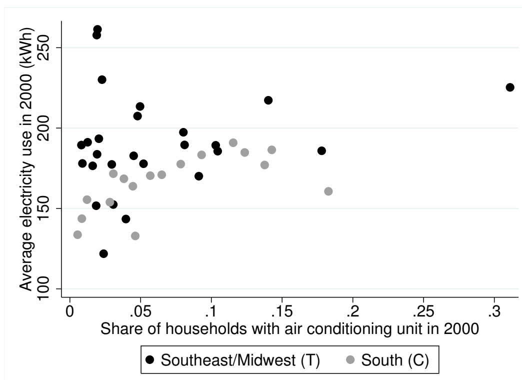
Figure G.3: Joint distribution of average electricity use and relevant covariates before the crisis III (2000)



Each observation corresponds to a distribution utility and its concession area. The panels display the joint distribution of monthly average residential electricity use per customer and relevant covariates before the crisis (2000) for distribution utilities in the Southeast/Midwest and in the South.

Figure G.4: Joint distribution of average electricity use and relevant covariates before the crisis IV (2000)

(a) Share with air conditioning

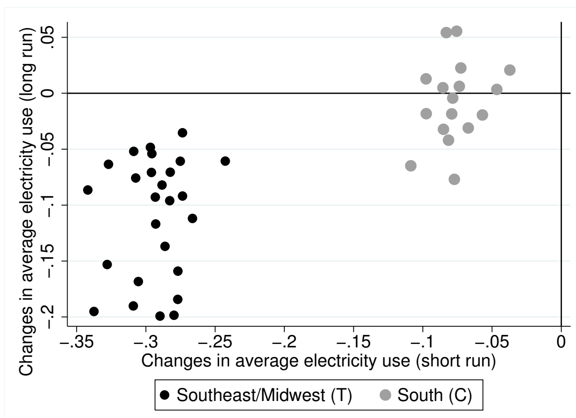


Each observation corresponds to a distribution utility and its concession area. The panels display the joint distribution of monthly average residential electricity use per customer and relevant covariates before the crisis (2000) for distribution utilities in the Southeast/Midwest and in the South.

H Changes in average electricity use during and after the crisis for distribution utilities in the Southeast/Midwest and in the South

Figure H.1 displays the changes in average residential electricity use per customer in every distribution utility in the Southeast/Midwest and in the South during the crisis (June 2001-February 2002) and long after the crisis (June 2010-February 2011) compared to before the crisis (June 2000-February 2001). Changes in average residential electricity use per customer y are calculated as follows: $\Delta y = \left(y_{during/after} - y_{before}\right)/y_{before}$. The figure shows that changes in average residential electricity use were lower for every distribution utility in the Southeast/Midwest during the crisis. The figure shows that changes in average residential electricity use were lower for almost every distribution utility in the Southeast/Midwest long after the crisis.

Figure H.1: Changes in average electricity use during and long after the crisis w.r.t before the crisis

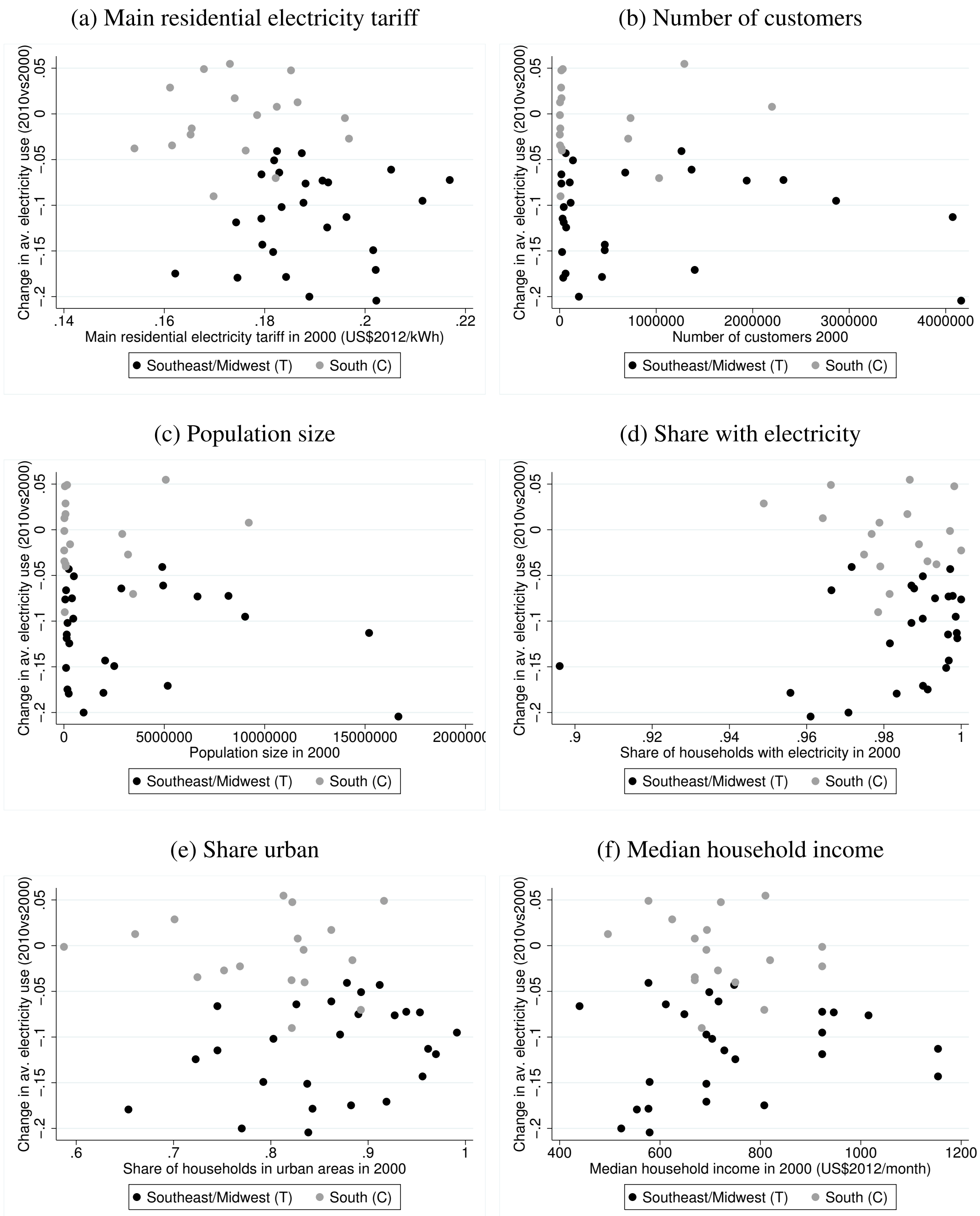


The figure displays the changes in average residential electricity use per customer in every distribution utility in the Southeast/Midwest and in the South during the crisis (June 2001-February 2002) and long after the crisis (June 2010-February 2011) compared to before the crisis (June 2000-February 2001). Changes in average residential electricity use per customer are calculated as follows: $\Delta y = \left(y_{during/after} - y_{before}\right)/y_{before}$.

I The joint distribution of long-run changes in average electricity use (2010 vs. 2000) and of levels of covariates in 2000 for distribution utilities in the Southeast/Midwest and in the South

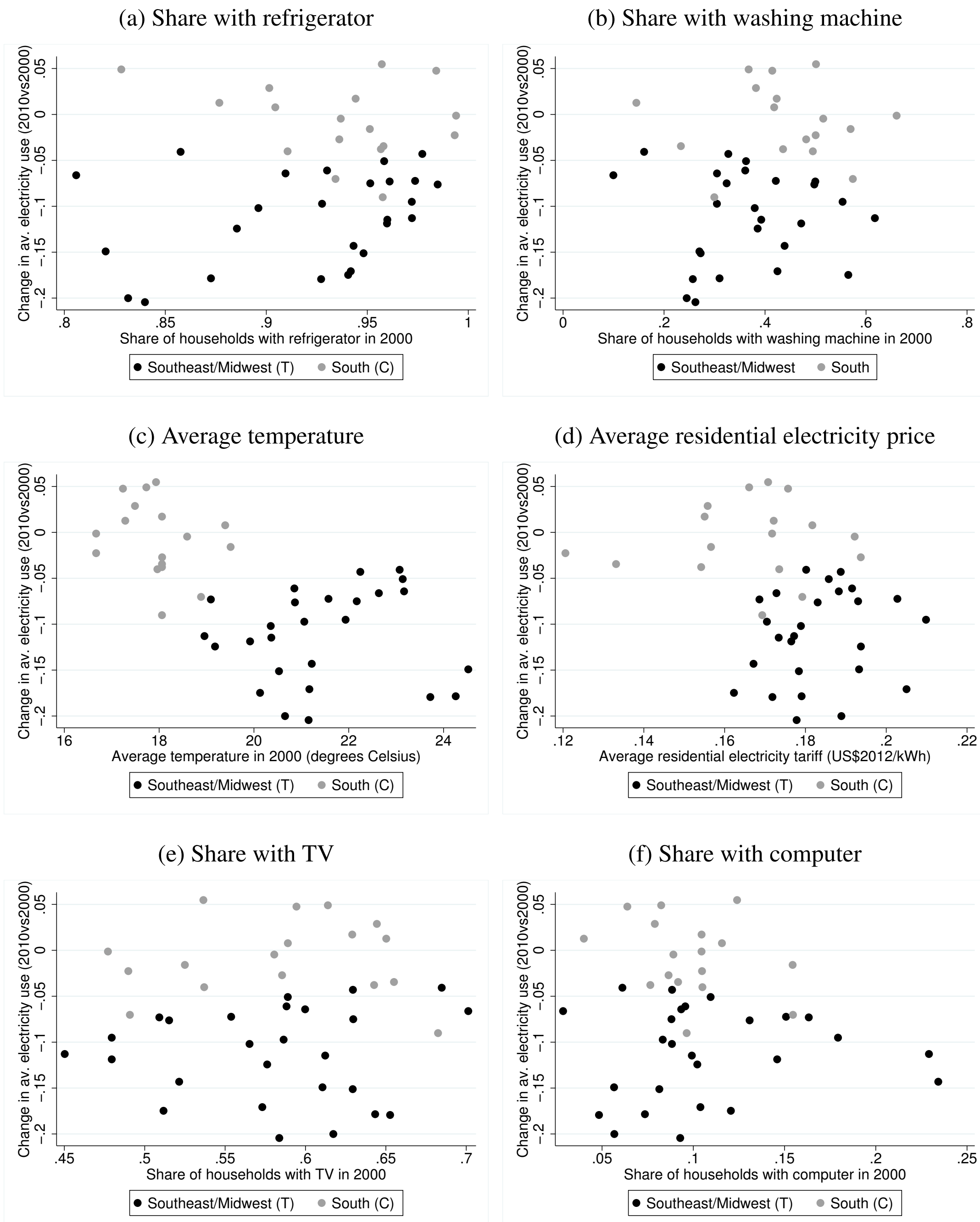
The following figures display the joint distribution of long-run changes in monthly average residential electricity use per customer between 2010 (after the crisis) and 2000 (before the crisis) and of levels of relevant covariates in 2000 for distribution utilities in the Southeast/Midwest and in the South. Long-run changes for a variable y are calculated as follows: $\Delta y = (y_{2010} - y_{2000})/y_{2000}$. We present figures for all the variables displayed in Tables 1 and A.1. The data come from either the utility-level data from ANEEL or from the 2000 and 2010 censuses matched to the concession area of each distribution utility. The figures show that long-run changes in average electricity use per customer are systematically lower in the Southeast/Midwest than in the South for given levels of those covariates at baseline.

Figure I.1: Long-run changes in average electricity use and levels of covariates at baseline I



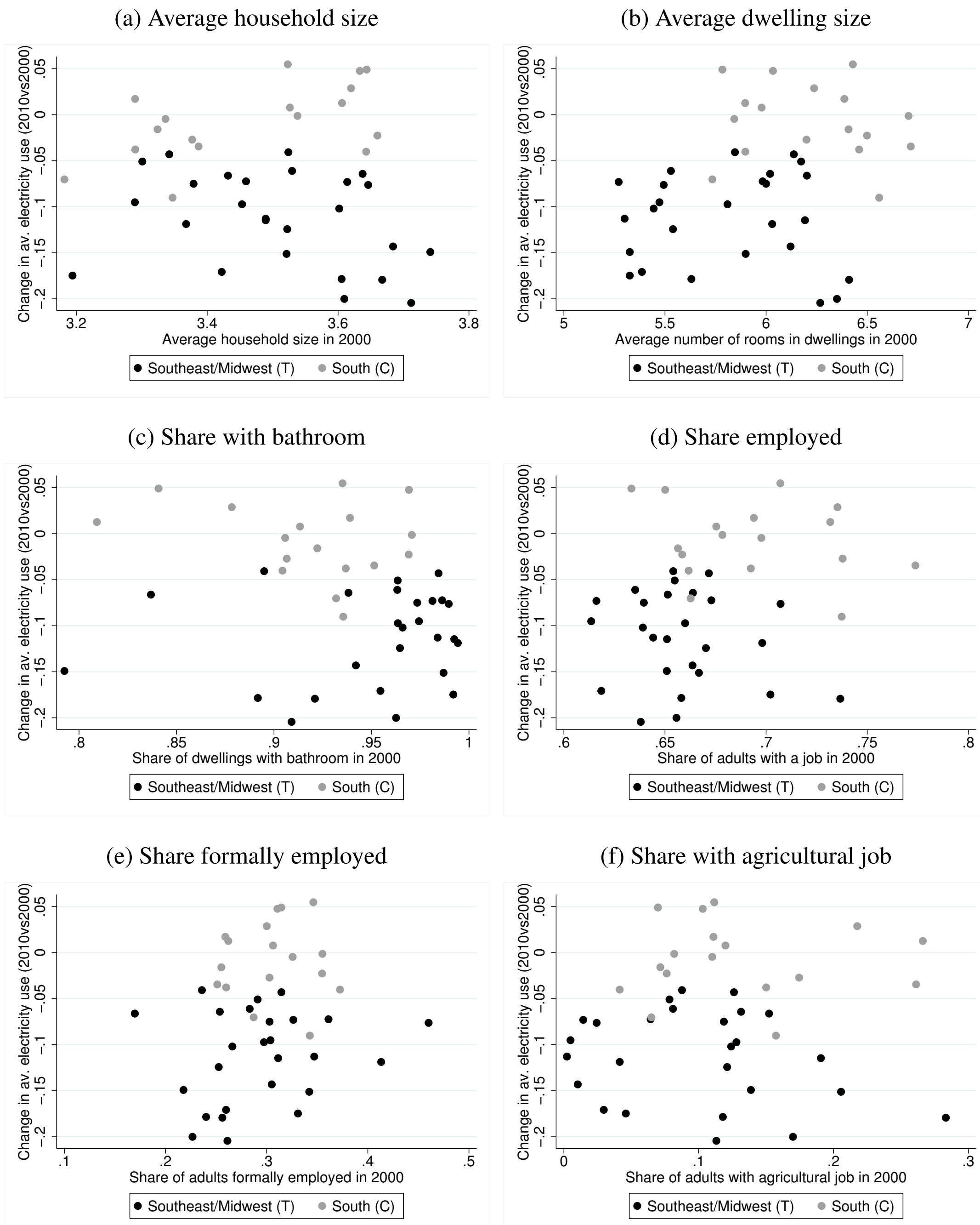
The panels display the joint distribution of long-run changes in average residential electricity use per customer between 2010 (after the crisis) and 2000 (before the crisis) and of levels of relevant covariates in 2000 for distribution utilities in the Southeast/Midwest and in the South. Long-run changes for a variable y are calculated as follows: $\Delta y = (y_{2010} - y_{2000})/y_{2000}$.

Figure I.2: Long-run changes in average electricity use and levels of covariates at baseline II



The panels display the joint distribution of long-run changes in average residential electricity use per customer between 2010 (after the crisis) and 2000 (before the crisis) and of levels of relevant covariates in 2000 for distribution utilities in the Southeast/Midwest and in the South. Long-run changes for a variable y are calculated as follows: $\Delta y = (y_{2010} - y_{2000})/y_{2000}$.

Figure I.3: Long-run changes in average electricity use and levels of covariates at baseline III



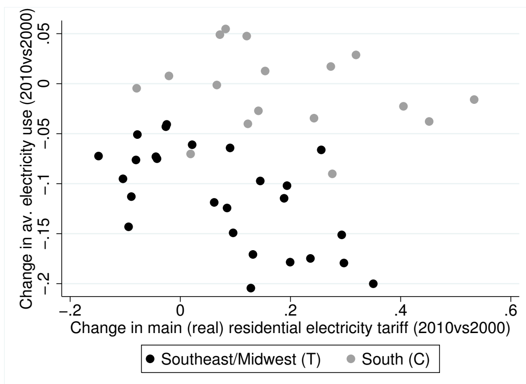
The panels display the joint distribution of long-run changes in average residential electricity use per customer between 2010 (after the crisis) and 2000 (before the crisis) and of levels of relevant covariates in 2000 for distribution utilities in the Southeast/Midwest and in the South. Long-run changes for a variable y are calculated as follows: $\Delta y = (y_{2010} - y_{2000})/y_{2000}$.

J The joint distribution of long-run changes in average electricity use and in covariates (2010 vs. 2000) for distribution utilities in the Southeast/Midwest and in the South

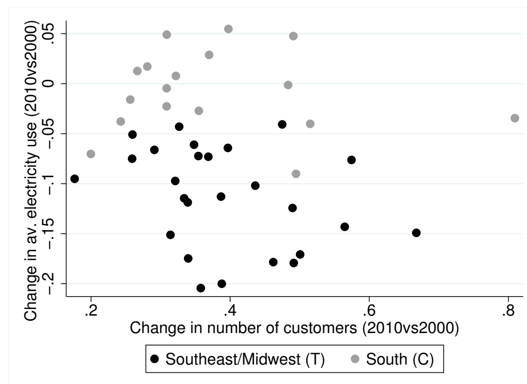
The following figures display the joint distribution of long-run changes in average residential electricity use per customer and in relevant covariates between 2010 (after the crisis) and 2000 (before the crisis) for distribution utilities in the Southeast/Midwest and in the South. Long-run changes for a variable y are calculated as follows: $\Delta y = (y_{2010} - y_{2000})/y_{2000}$. We present figures for all the variables displayed in Tables 1 and A.1. The data come from either the utility-level data from ANEEL or from the 2000 and 2010 censuses matched to the concession area of each distribution utility. The figures show that there is some overlap in the distributions of long-run changes in all covariates in the Southeast/Midwest and in the South. They also show that there is a lot of variation in terms of long-run changes in those covariates among distribution utilities. Importantly, the figures show that long-run changes in average electricity use per customer are systematically lower in the Southeast/Midwest than in the South for given long-run changes in those covariates. This explains why our results are robust to controlling for relevant covariates in Table A.2.

Figure J.1: Long-run changes in average electricity use and in relevant covariates I

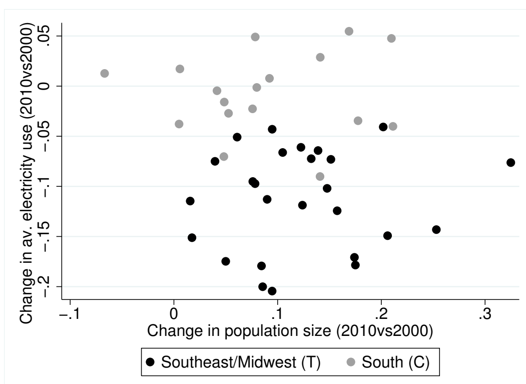
(a) Changes in main residential electricity tariff



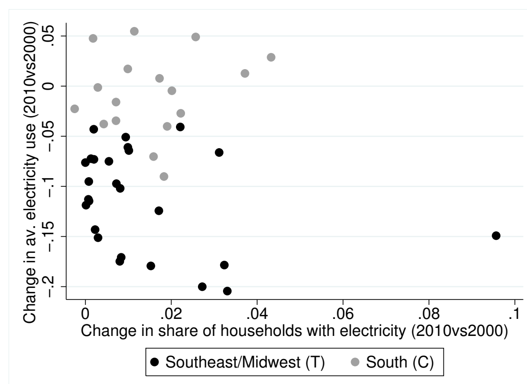
(b) Changes in number of customers



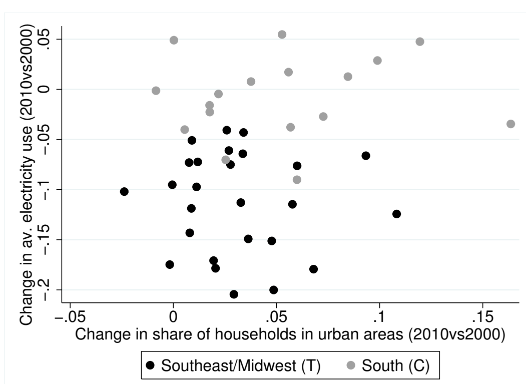
(c) Changes in population size



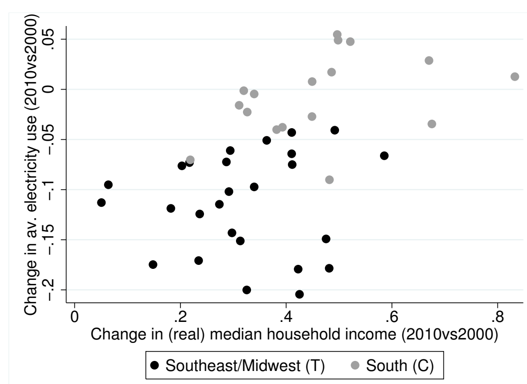
(d) Changes in share with electricity



(e) Changes in share urban



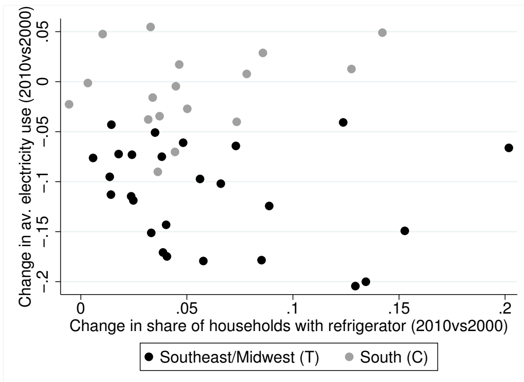
(f) Changes in median household income



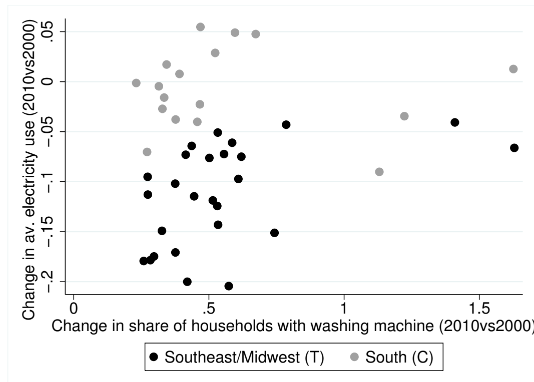
The panels display the joint distribution of long-run changes in average residential electricity use per customer and in relevant covariates between 2010 (after the crisis) and 2000 (before the crisis) for distribution utilities in the Southeast/Midwest and in the South. Long-run changes for a variable y are calculated as follows: $\Delta y = (y_{2010} - y_{2000})/y_{2000}$.

Figure J.2: Long-run changes in average electricity use and in relevant covariates II

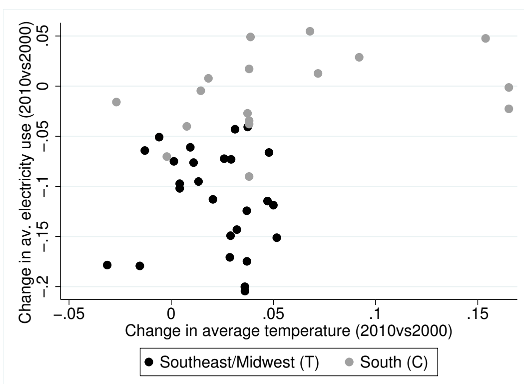
(a) Changes in share with refrigerator



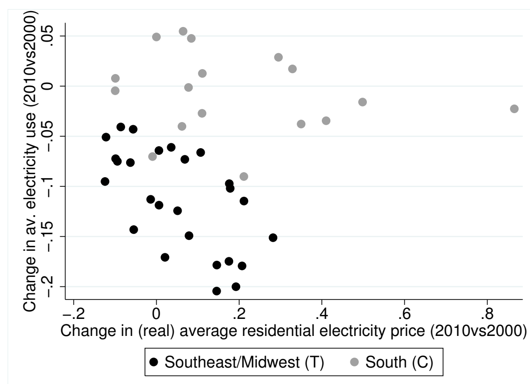
(b) Changes in share with washing machine



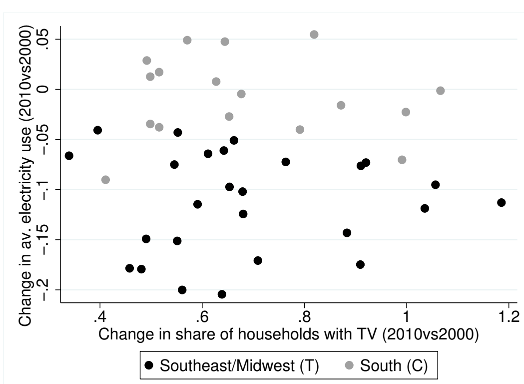
(c) Changes in average temperature



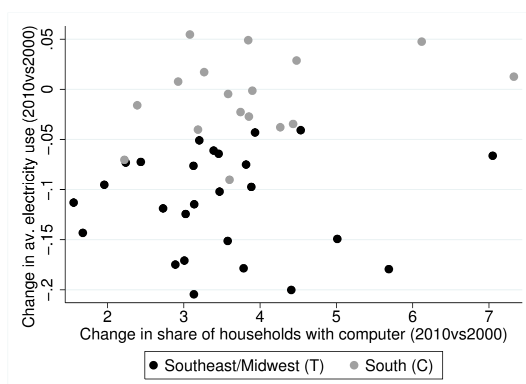
(d) Changes in average residential electricity price



(e) Changes in share with TV

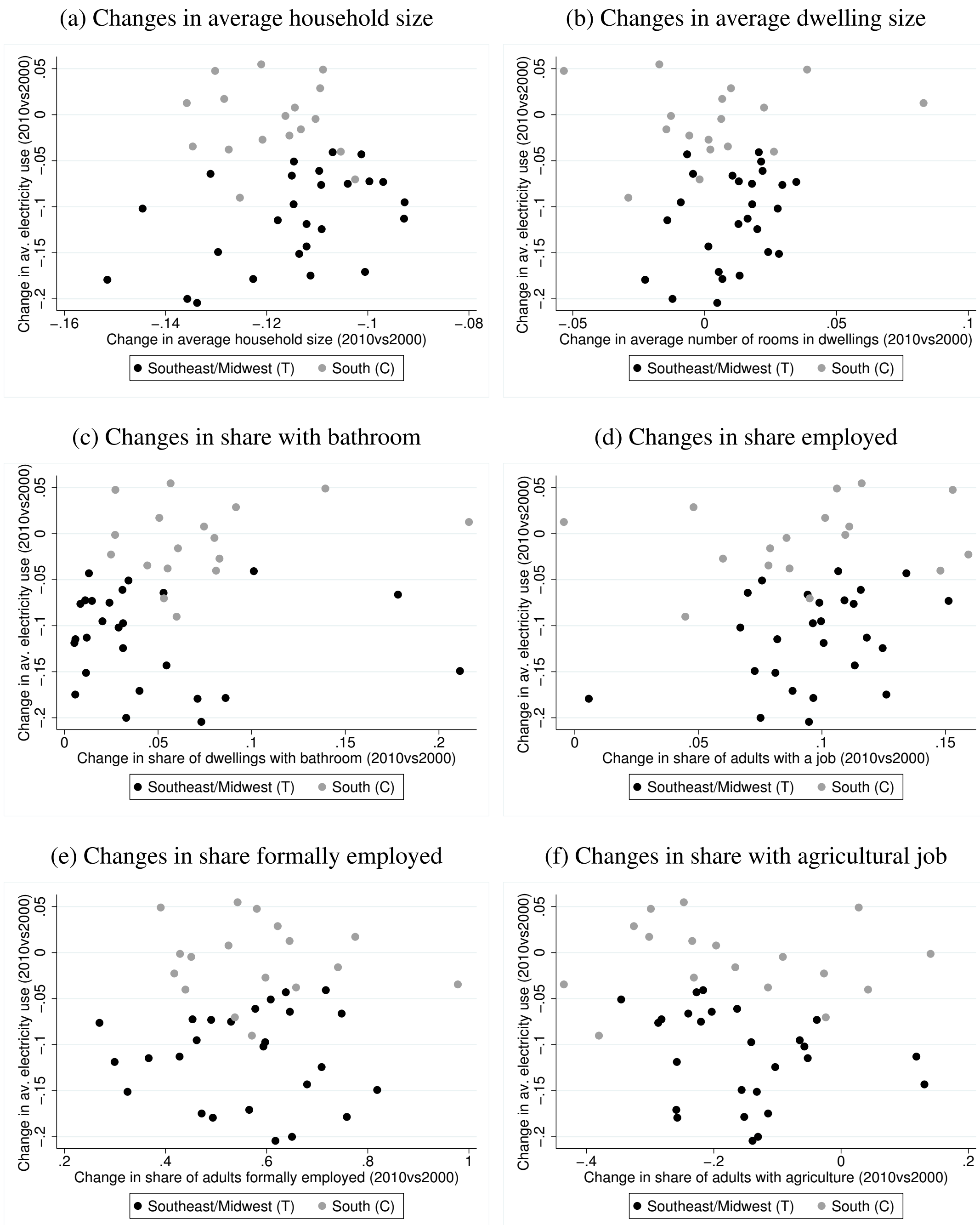


(f) Changes in share with computer



The panels display the joint distribution of long-run changes in average residential electricity use per customer and in relevant covariates between 2010 (after the crisis) and 2000 (before the crisis) for distribution utilities in the Southeast/Midwest and in the South. Long-run changes for a variable y are calculated as follows: $\Delta y = (y_{2010} - y_{2000})/y_{2000}$.

Figure J.3: Long-run changes in average electricity use and in relevant covariates III



The panels display the joint distribution of long-run changes in average residential electricity use per customer and in relevant covariates between 2010 (after the crisis) and 2000 (before the crisis) for distribution utilities in the Southeast/Midwest and in the South. Long-run changes for a variable y are calculated as follows: $\Delta y = (y_{2010} - y_{2000})/y_{2000}$.

K Electricity theft (or non-technical loss)

This section presents the additional analysis on electricity theft mentioned in the paper. Although electricity theft is a big issue for distribution utilities, there is no good data on electricity theft in Brazil. Here, we investigate if the energy crisis affected electricity theft differentially in treated and control distribution utilities using a difference-in-differences specification and two different datasets: (i) reported data on distribution losses from ANEEL; and (ii) household self-reported data in the Household Expenditure Surveys (POF). We find no robust evidence that electricity theft was affected by the energy-saving program.

K.1 Difference-in-differences estimates using data on distribution losses

First, we use yearly reports for 24 utilities in the Southeast/Midwest (13) and in the South (11) from 1998 to 2008.

Distribution utilities are supposed to report yearly information on distribution losses, which correspond to the electricity load not charged to particular customers, to the regulator, but many distribution utilities did not provide this information prior to 2000. ANEEL provided the available yearly data on distribution losses by distribution utility. The data include six fields: name of the utility, year, electricity load, total distribution losses, "technical" distribution losses, "non-technical" distribution losses. Distribution losses are divided into technical (engineering estimates) and non–technical (residual) losses. "Non-technical" losses are supposed to capture electricity theft (distribution losses unexplained by engineering estimates). It is unclear how distribution utility separately identify the two categories and the resulting information is noisy.

Table K.1 displays coefficients from regressing several outcomes (listed above each column) on year dummies interacted with an indicator for utilities subject to the energy-saving program during the crisis (difference–in–differences estimators in every year). The reference year corresponds to 2000. Regressions include uninteracted year dummies and distribution utility fixed effects, and control for the main electricity tariff, total population, total formal employment, and average temperature (log) for each distribution utility. Column (1) considers a specification similar to our main difference-in-differences specification in the paper but looks at total residential consumption at the yearly level for this sample of distribution utilities. The long-term effect on average residential electricity use is very similar. Columns (2)–(5) use the data from the yearly reports on distribution losses. Note that those data are for the whole distribution utility, and not specific to its residential customers. The outcome in column (2) is the total load reported by the distribution utility; the outcome in column (3) is the total losses reported by the distribution utility; the outcome in column (4) is the total *technical* losses reported by the distribution utility; the outcome in column (5) is the total *non–technical* losses reported by the distribution utility. As expected, the total load decreased during and after the crisis. The effects are large and include other types of customers (e.g., industrial), so the long-term effects might include changes in the industrial composition of firms served by particular distribution utilities. Total losses were also reduced, which is not surprising if they are proportional to the total load, but the data are noisy, and therefore our estimates are not significant during and right after the crisis. Once we divide total losses into technical and non–technical losses, we find some evidence that technical losses decreased, although estimates are again noisy. We find no evidence that non-technical losses increased, which would be the case if electricity theft increased and non–technical losses were a good proxy for theft. The data are very noisy. For instance, point estimates imply that non–technical losses increased by .268 log points from 1999 to 2000 in the Southeast/Midwest compared to the South and then decreased by .243 log points the following year.

K.2 Difference-in-differences estimates using household expenditure surveys

Second, we use a proxy for household-level electricity theft based on the Brazilian Household Expenditure Surveys (POF, with rounds in 1996-1997, 2002-2003, and 2008-2009).

The surveys, which contain household-level microdata (repeated cross-sections), were conducted by the Brazilian Geography & Statistics Institute (IBGE), which is also responsible for the National Census. All surveys were conducted between July of the base year and June of the following year. The 1996/1997 survey was based on a representative sample of the population of the largest metropolitan areas of Brazil. The two subsequent surveys were based on a representative sample of the whole Brazilian population. The microdata is not available at the municipality level and cannot be matched to the concession areas of distribution utilities. In our analysis, we restrict attention to the metropolitan areas of the seven states in the Southeast/Midwest and in the South, for which the samples are representative in all survey rounds.

There is a relevant difference between the sampling of POF surveys and the sampling of the datasets used so far. The administrative data from ANEEL, the billing data from LIGHT, and the PPH surveys capture only households regularly and legally connected to the electricity grid. POF surveys, however, aim to be representative of all households including those who have no regular or legal connection to electricity. Consequently, some households in POF surveys own electrical appliances, but claim to have no expenses for electricity and to not own a generator. We use this information to investigate the share of households who are likely illegally connected to the electricity grid (electricity theft).

Our outcome of interest is a dummy variable equal to 1 if a household reports paying for electricity and equal to 0 if a household reports not paying for electricity (and has at least one electrical appliance) or owning an electricity generator.[86] We then estimate a difference-in-differences specification analogous to the one used in Section 3:

$$Y_{h,d,t} = \alpha_d + \sum_{t' \in \{2002,2008\}} \left\{ \delta_{t'} \, \mathbb{1}(t = t' \,\&\, d \in \text{SE/MW}) \right\} + \gamma_t + X_{h,d,t} + \nu_{h,d,t} \tag{75}$$

where $Y_{h,d,t}$ is an outcome for household h from state d in survey round t. We control for state fixed effects α_d and a survey round fixed effect γ_t. The coefficients δ_{2002} and δ_{2008} are the short- and long-run difference-in-differences estimators under a common-trend assumption. We control for household characteristics, $X_{h,d,t}$, which may be correlated with different trends in appliance ownership. The vector of household characteristics includes income, income squared, number of household members, and number of rooms.

Table K.2 shows that our point estimates are close to 0 both in the short-run and in the long run.

Table K.2: Difference-in-differences results for the share of households paying for electricity

	Paying for Electricity
	(1)
$SE/MW \times Year2002$	**-.002**
	(.044)
$SE/MW \times Year2008$	**.009**
	(.025)
Average SE/MW 1996	.901
N	34493

The table uses household-level microdata for seven states (urban area only) in the South and Southeast/Midwest subsystems from the Household Expenditure Surveys (POF) conducted in 1996-1997, 2002-2003 and 2008-2009. The table displays difference-in-differences estimates for the impact of the energy-saving program on the share of households that pay for electricity based on the specification in equation (77). The dependent variable is a dummy variable equal to 1 if a household reports paying for electricity and equal to 0 if a household reports not paying for electricity (and has at least one electrical appliance) or owning an electricity generator. We combine electricity theft and own electricity generation because the consequences are the same for our results: if households responded by increasing either one of these two behaviors, our estimates in Section 3 would overestimate the impact of the energy-saving program. All regressions include distribution utility fixed effects and year fixed effects, and control for income, income squared, number of household members, and number of rooms. We input missing values for income with a linear regression (separately for each year) of income on the other controls. Significance levels: *10%, **5%, ***1% (s.e. clustered by state, using the wild-cluster bootstrap).

[86] We combine electricity theft and own electricity generation because the consequences are the same for our results: if households responded by increasing either one of these two behaviors, our estimates in Section 3 would overestimate the impact of the energy-saving program.

Table K.1: Difference–in–differences results for reported distribution losses

	Residential electricity use (logs)		Total load (logs)		Total losses (logs)		Technical losses (logs)		Non–technical losses (logs)	
	(1)		(2)		(3)		(4)		(5)	
1998	.0192	(.0143)	.03	(.0214)	-.081	(.0978)	-.0288	(.1008)	-.2451	(.2126)
1999	.0151	(.0125)	.0434**	(.0199)	-.0264	(.0947)	.0597	(.1214)	-.268**	(.1301)
2001	-.1194***	(.0189)	-.1362***	(.0235)	-.1725	(.2097)	-.1377	(.2044)	-.2427	(.3722)
2002	-.1724***	(.0091)	-.1647***	(.0198)	-.1422	(.119)	-.1597	(.1135)	-.2059	(.2473)
2003	-.1352***	(.0135)	-.1277***	(.0222)	-.148	(.093)	-.1829**	(.0836)	-.2198	(.2366)
2004	-.1421***	(.015)	-.1496***	(.0243)	-.1001	(.1596)	-.1167	(.1435)	-.2188	(.3258)
2005	-.1213***	(.0208)	-.154***	(.0342)	-.2131*	(.1243)	-.1937	(.1302)	-.3477	(.2367)
2006	-.1241***	(.0224)	-.1763***	(.0424)	-.2152**	(.091)	-.1776*	(.0918)	-.3932	(.2402)
2007	-.1465***	(.0205)	-.2062***	(.0457)	-.2289**	(.093)	-.2201**	(.0922)	-.4094**	(.1887)
2008	-.1259***	(.0211)	-.1881***	(.0399)	-.2873***	(.0967)	-.2589***	(.0903)	-.4276**	(.2179)
Log main tariff	-.1524***	(.0541)	-.0608	(.0592)	-.2439	(.1977)	-.2348	(.1555)	-.2325	(.3227)
Log population	.3926**	(.1744)	.6961**	(.3189)	.3423	(.7843)	-.241	(.6104)	2.897	(1.917)
Log formal employment	.3482***	(.0781)	.7891***	(.1748)	.2796	(.2944)	.7233***	(.2005)	-.5474	(.5975)
Log average temperature	.565**	(.223)	.4974*	(.2805)	1.312	(1.547)	1.681	(1.494)	2.342	(2.91)
Clusters	24		24		24		24		24	
Observations	3168		3168		3168		3168		3168	

Significance levels: *10%, **5%, ***1% (s.e. clustered by distribution utility in parentheses). Data for 24 distribution utilities in the Southeast/Midwest (13) and in the South (11) reporting technical and non–technical losses prior to 2000. Yearly data from 1998 to 2008. Distribution losses are the share of the load not charged to particular customers. Distribution losses are divided into technical (engineering estimates) and non–technical (residual, a noisy proxy of theft) losses. It is unclear how companies separately identify the two categories; the resulting information is noisy. The table displays coefficients from regressing several outcomes (listed above each column) on year dummies interacted with an indicator for utilities subject to the energy-saving program during the crisis (difference–in–differences estimators in every year). The reference year corresponds to 2000. Regressions include uninteracted year dummies and utility fixed effects. Column (1) uses a specification similar to our main difference-in-differences results in Section 3 but looks at total residential consumption at the yearly level for this sample of utilities. Columns (2)–(5) use the data from the yearly reports on distribution losses.

L Robustness of the patterns based on individual billing data shown in Figure 4

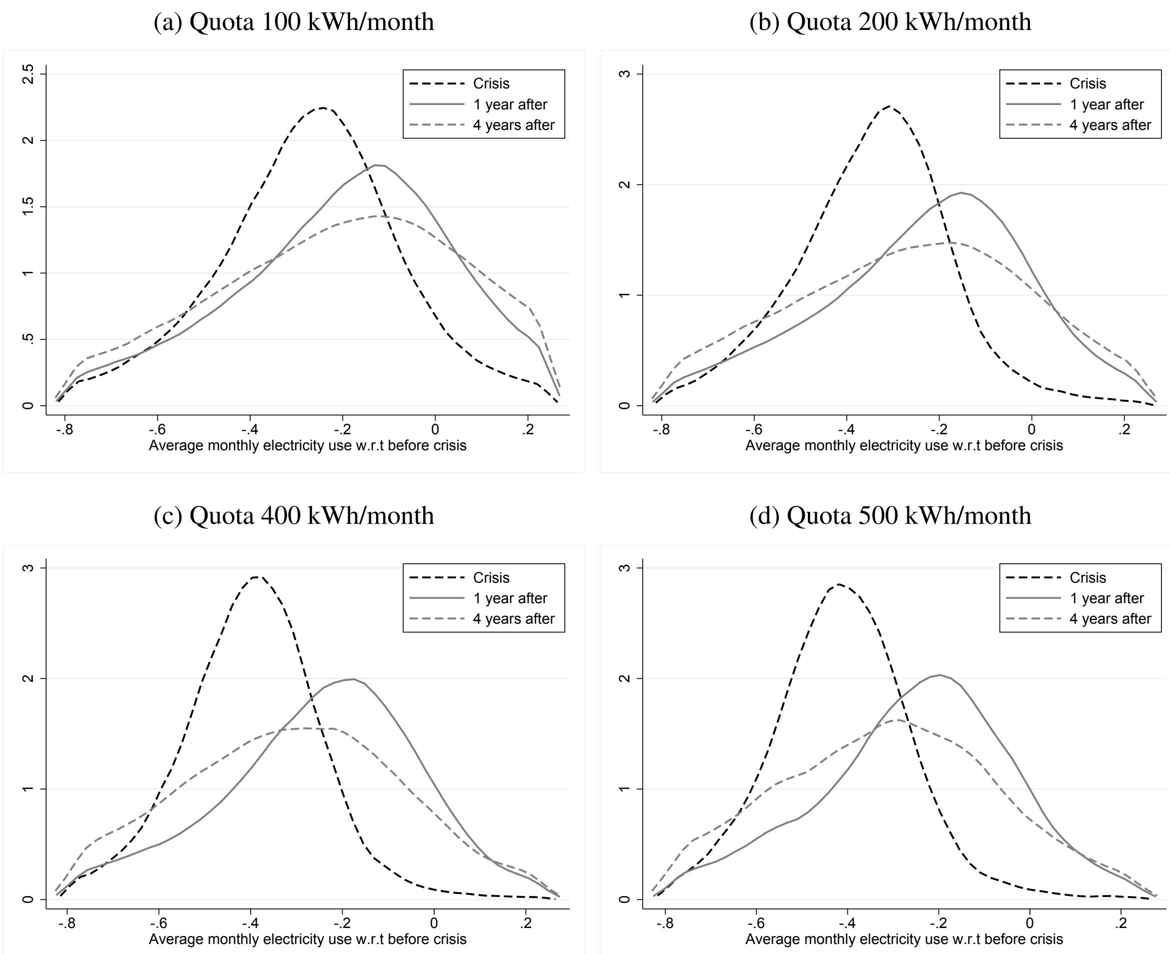
This section shows that we find the same patterns based on the household-level billing data as in Figures 5c and 5d in Section 4.1 when we use different samples of Light customers.

Figure L.1 displays the distribution of changes in average electricity use during and after the crisis compared to the same months (July to December) before the crisis, as in Figure 5c. We use the balanced panel of customers and consider customers who had about the same initial quota and thus about the same baseline consumption such that they faced the same incentives during the crisis, as in Figure 5c. However, we consider different quota levels. In particular, we consider customers with quotas 10% above and below 100 kWh/month (253,748 customers), 200 kWh/month (155,007 customers), 400 kWh/month (41,137 customers), and 500 kWh/month (21,637 customers).

During the crisis, 87%, 96%, 98%, and 98% reduced electricity use and the median customer reduced usage by 26%, 33%, 40%, and 41% among customers with baseline levels around 100 kWh/month, 200 kWh/month, 400 kWh/month, and 500 kWh/month, respectively. Four years after the crisis, 62%, 79%, 87%, and 86% were still using less electricity than before the crisis; the median customer was using 13%, 24%, 31%, and 31% less electricity among customers with baseline levels around 100 kWh/month, 200 kWh/month, 400 kWh/month, and 500 kWh/month, respectively. Changes in consumption levels were thus large at every baseline consumption level, but they were larger for customers with higher baseline levels.

Figure L.2 displays the correlation between individual changes in electricity use during the crisis and four years after the crisis compared to the same months before the crisis, as in Figure 5d, but for the same samples as in Figure L.1. Customers are averaged by bins of 5% changes in electricity use during the crisis. The strong correlations provide household-level evidence that changes in electricity use that took place during the crisis were very persistent at every baseline level.

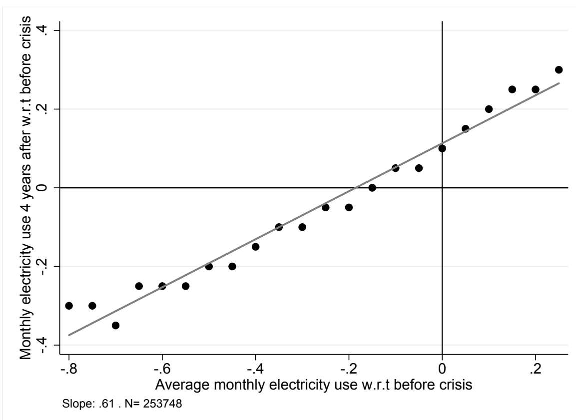
Figure L.1: Distribution of changes in monthly electricity use for balanced panels of customers with the same baseline level and thus the same quota during the crisis (robustness of Figure 5c)



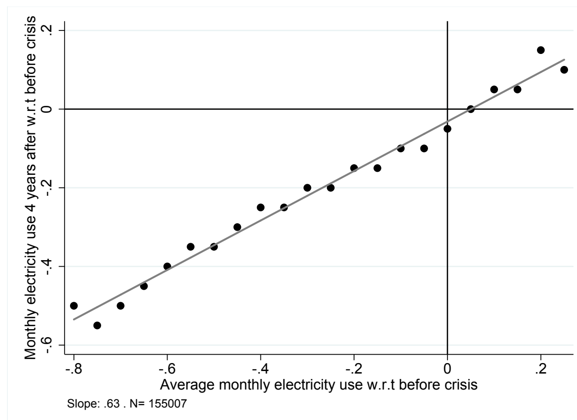
The panels display the distribution of changes in average electricity use during and after the crisis compared to the same months before the crisis, for subsets of the same balanced panel of customers. The panels consider customers who had about the same baseline for quota assignment (and thus the same quota) such that they faced the same incentives during the crisis, as in Figure 5c in the paper. However, we consider different quota levels. In particular, we consider customers with quota levels 10% above and below 100 kWh/month (252,820 customers), 200 kWh/month (154,917 customers), 400 kWh/month (41,108 customers), and 500 kWh/month (21,585 customers). Kernel densities use Epanechnikov kernels and optimal bandwidths.

Figure L.2: Correlation between changes in electricity use during the crisis and four years after the crisis for balanced panels of customers with same baseline level and thus the same quota (robustness of Figure 5d)

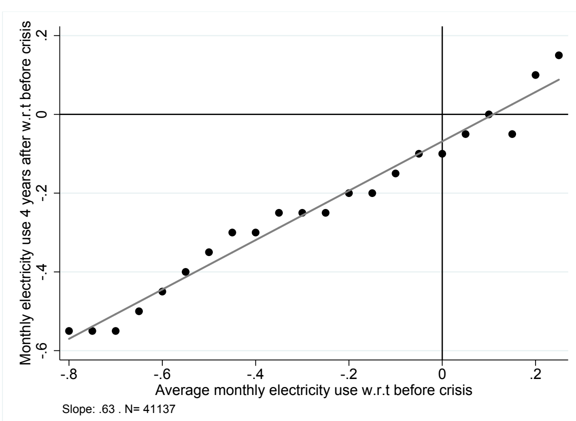
(a) Quota 100 kWh/month



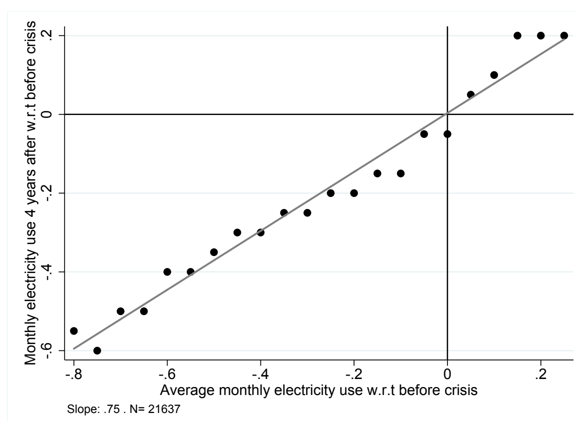
(b) Quota 200 kWh/month



(c) Quota 400 kWh/month



(d) Quota 500 kWh/month



The panels display the correlation between individual changes in electricity use during the crisis and four years after the crisis compared to the same months before the crisis, as in Figure 5d in the paper, but for the same samples as in Figure L.1. Customers are averaged by bins of 5% changes in electricity use during the crisis.

M Threat of power cuts: spatial regression discontinuity (RD)

This section investigates the potential contribution of the threat of power cuts to the observed change in households' electricity use during the crisis. According to the energy-saving program, customers who exceeded their quotas repeatedly would have their power cut for a period of three to six days. This was designed as an extreme incentive scheme to induce consumption change. However, distribution utilities did not have enough staff to implement power cuts,[87] so power cuts were in practice limited to customers who repeatedly consumed far above their quota. Power cuts were even prohibited in some areas. For instance, this was the case for the city of Rio de Janeiro (Lei Municipal 3266/2001), one of the cities supplied by LIGHT.

We implement a spatial regression discontinuity design to look at differential response to the energy-saving program by LIGHT customers around the border between the city of Rio de Janeiro and neighboring municipalities (the administrative concepts of city and municipality are equivalent in Brazil) served by LIGHT. If the threats of power cuts were effective at inducing changes in electricity use, we should expect greater reductions in electricity use by customers subject to the power cuts (those outside the city of Rio de Janeiro) than by those not facing these threats (those within the city of Rio de Janeiro).

Figure M.1 displays a map of the northern part of the city of Rio de Janeiro with the city border in red (solid line). This is the only part of the city where a land border separates the city of Rio de Janeiro and other municipalities (the southern and western parts are bordered by the ocean; the eastern part, shown on the map, is the Guanabara Bay with its islands). The background color gradient captures the distance to the border in shades of gray (the darker, the more distant from the border). In the billing data, we have 3340 zip codes with LIGHT customers, which we georeferenced using GoogleMaps (package 'ggmap' in R). The red dots show the centroids of the zip codes in the map. The areas with few red dots correspond to the many forested hills in the region.

For the analysis, we use the balanced panel of customers observed in each month from 2000 to 2005. Moreover, we restrict attention to customers whose consumption level was above the median prior to the crisis (the power cuts had to be targeted to large customers repeatedly exceeding their quota). As before, we also restrict attention to the months of July to December in each year.

Figure M.2 shows the average electricity use per customer in 2000 (panel a) and 2001 (panel b) and the average change between 2000 and 2001 panel (c) by 40 equal-sized bins of distances from the border of the city of Rio de Janeiro city, up to 5 kilometers away from the border. A positive value for the running variable indicates that the household resides within the border of the

[87] See, e.g., https://www.correiodobrasil.com.br/light-nao-tem-capacidade-para-cortar-luz-de-quem-excedeu-a-taxa-de-consumo/

city of Rio de Janeiro. The vertical bars depict 95% confidence intervals for the average within each bin. The red line shows a linear fit on each side of the border, where we weight each bin by the number of household it contains. We see that average consumption levels in 2000 (panel a) and 2001 (panel b) seem to be smaller on the Rio de Janeiro's side of the border. However, we find no discernible difference at the border in terms of consumption changes during the crisis (panel c).

The patterns in Figure M.2 show that a typical RD design would unlikely be valid in our case, as there seems to be systematic differences at baseline between households on either side of the border. However, we can use a difference-in-discontinuity approach – looking at differential changes over time at the border – to identify the causal effect of the threat of power cuts under a common-trend assumption for units on either side of the border.

Table M.1 displays the results from estimating the following specification:

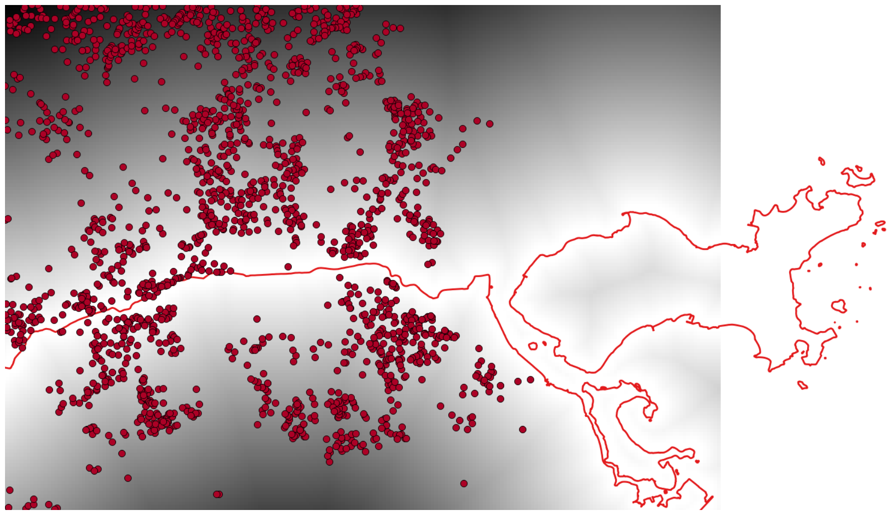
$$Y_i = \alpha + \gamma \cdot RioCity_i + f\left(DistBorder_i\right) + \varepsilon_i \tag{76}$$

where Y_i is in turn the average consumption level in 2000, the average consumption level in 2001, and the change in consumption in 2001 relative to 2000 (in logs), for household i. The variable $RioCity_i$ is a dummy variable, which is equal to one if household i resides in the city of Rio de Janeiro. The variable $DistBorder_i$ is the distance between the border and the centroid of the zip code where household i resides. We specify the function $f\left(DistBorder_i\right)$ as $f\left(DistBorder_i\right) = Rio_i * f^{Rio}\left(DistBorder_i\right) + \left(1 - Rio_i\right) * f^{OutsideRio}\left(DistBorder_i\right)$. The function $f\left(DistBorder_i\right)$ is thus a polynomial in distance from the border that is allowed to differ on either side of the border. In practice, we use a linear polynomial on each side of the border, a rectangular kernel, and the optimal bandwidth according to Imbens and Kalyanaraman (2012), which is different for each outcome. Finally, we cluster the error term ε_i by zip code to allow for some spatial correlation.

The estimate of interest, which we display in Table M.1, is $\widehat{\gamma}$. It captures the difference in average consumption at the border before and during the crisis for the outcomes in columns (1) and (2), respectively. It captures a difference-in-discontinuity estimate for the impact of the threat of power cuts for the outcome in column (3), the change in consumption from 2000 to 2001.

Results in Table M.1 are consistent with the graphical evidence in Figure M.2. We estimate that average consumption is about 12.4%-12.5% lower at the border on the side of the city of Rio de Janeiro in columns (1) and (2). The difference is also statistically significant. Moreover, if anything, it decreased from 2000 to 2001, even though households in the city of Rio de Janeiro were not subject to any power cut for exceeding their quota. Accordingly, we find no differential change over time on either side of the border in column (3). The point estimate is positive but close to 0 and insignificant (t-statistic = 0.61; note that the optimal bandwidth is larger for this outcome, at 4.6km). This suggests either (i) that the threat of power cut did not directly influence households' electricity use during the crisis, or (ii) that households were misinformed and did not changed consumption in response to the prohibition of power cuts in Rio de Janeiro.

Figure M.1: Map of the area around the border of the city of Rio de Janeiro



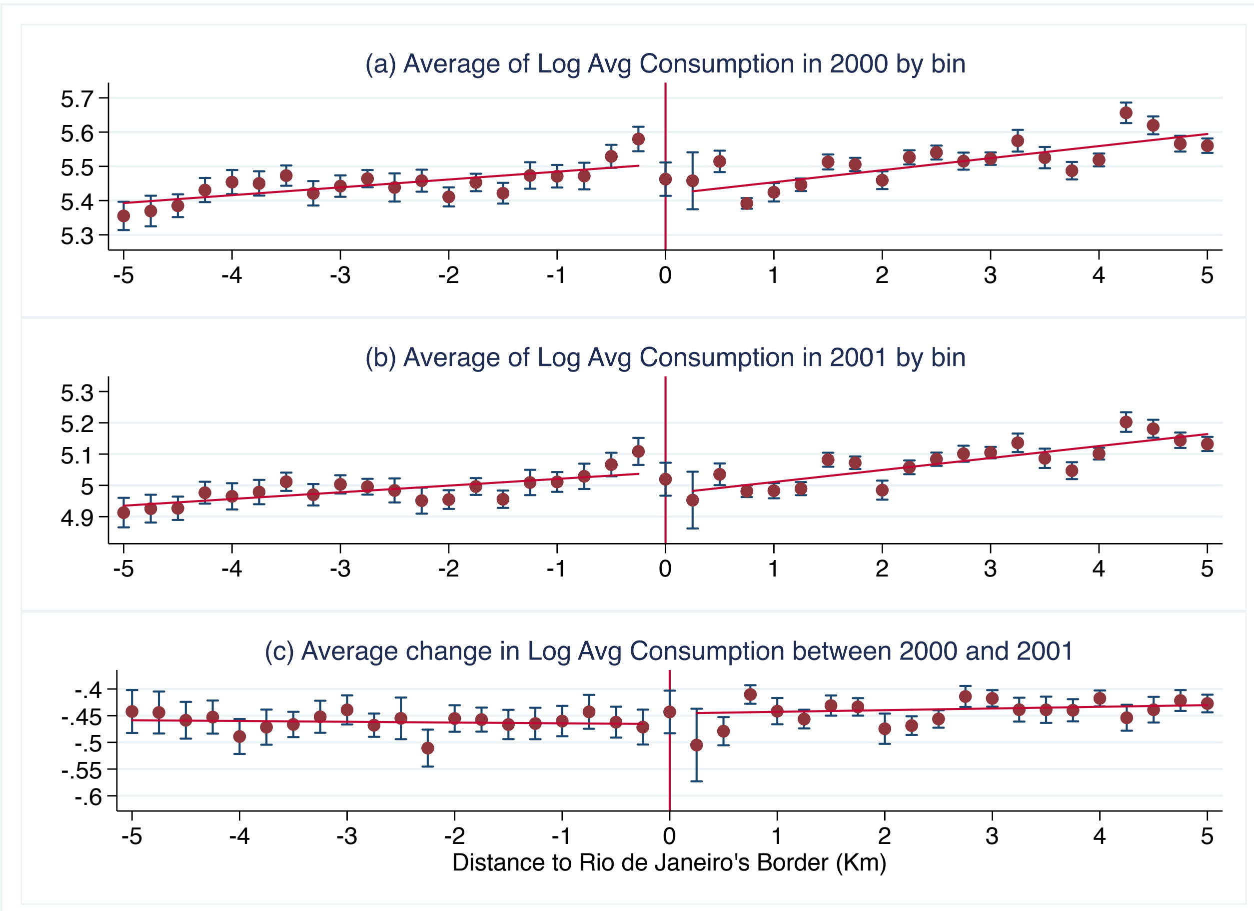
The figure displays a map of the northern part of the city of Rio de Janeiro with the city border in red (solid line). This is the only part of the city where a land border separates the city of Rio de Janeiro and other municipalities (the southern and western parts of the city are bordered by water; the eastern part, shown on the map, is the Guanabara Bay with its islands). The background color gradient captures the distance to the border in shades of gray (the darker, the more distant from the border). In the billing data, we have 3340 zip codes with LIGHT customers, which we georeferenced using GoogleMaps (package 'ggmap' in R). The red dots show the centroids of the zip codes in the map. The areas with few red dots correspond to the many forested hills in the region.

Table M.1: Results for the Regression Discontinuity around the border of the city of Rio de Janeiro

	Log Avg Consumption		Change in Log Avg Consumption
	2000	2001	2001–2000
	(1)	(2)	(3)
Rio City (γ)	-.124*	-.125**	.014
	(.067)	(.052)	(.023)
Optimal bandwidth (km)	1.3	1.5	4.6

This table shows the results from a Regression Discontinuity around the border of the city of Rio de Janeiro using a sub-sample of the balanced panel of Light customers observed regularly from 2000 to 2005 (see text). The running variable is the distance between the border and the centroid of the zip code where household i resides. The outcome variables in columns (1) and (2) are the log average monthly consumption between July and December in 2000 and 2001, respectively. The outcome in column (3) is the difference between 2001 and 2000 in the the log average monthly consumption between July and December in each year. We follow the specification in equation (76), using a linear polynomial on each side of the border, a rectangular kernel, and the optimal bandwidth according to Imbens and Kalyanaraman (2012), which is different for each outcome. Finally, we cluster the error term ε_i by zip code to allow for some spatial correlation. Number of customers in a 11km window around the border: 31,268 on the side of the city of Rio de Janeiro; 22,723 on the other side. Significance levels: $*10\%, **5\%, ***1\%$.

Figure M.2: Average consumption and average consumption changes before and during the crisis by distance to the border of the city of Rio de Janeiro



This figure shows the average household consumption in 2000 (a) and 2001 (b) and the average consumption change between 2000 and 2001 (c) by 80 equal-sized bins of distances from Rio de Janeiro city border, up to 5 kilometers away from the border. Positive distance represent households in Rio, while negative distance represent outside Rio. We restrict attention to consumption between July and December in each year. Sample of balanced panel with consumption above the median pre-crisis consumption. The vertical bars (not always visible) depict 95% confidence intervals of the local average within each bin. The red line shows the linear function of distance weighted by the number of observations in each bin.

N Additional evidence on mechanisms of hysteresis: PPH surveys

In the remaining sections, we present additional evidence on the mechanisms of hysteresis.

We can categorize mechanisms of hysteresis in three groups in our context. Households may have persistently changed the quantity of appliances that they owned, the type of appliances that they owned, or their utilization of these appliances. In this section, we shed some "quantitative" light on each of these mechanisms using the household-level microdata from the two most recent rounds of the PPH surveys. However, given the limited number of distribution utilities for which we have data, the exercise in this section is underpowered and so should be viewed as suggestive.

N.1 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 paper:

$$Y_{h,d,t} = \alpha_d + \gamma\, \mathbb{1}\,(t = 2005) + \delta\, \mathbb{1}(t = 2005\ \&\ d \in \text{SE/MW}) + log\left(X_{h,d,t}\right) + \nu_{h,d,t} \qquad (77)$$

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 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.[88] 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 Panel A of Table N.1 for the five main domestic appliances in terms of electricity use, and in Panel A of Table N.2 for other appliances. Standard errors are clustered by distribution utility using the wild cluster bootstrap-t (Cameron, Gelbach and Miller, 2008). The resulting confidence intervals are large given the small number of clusters, and typically include 0, so our results 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

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

with the information reported in Table 4. 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). The PPH survey measures instead the number of TVs per household.[89]

N.2 Appliances' characteristics

The PPH surveys recorded some appliance characteristics correlated with electricity use. We use the specification in equation (77) 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.

We display difference-in-difference estimates in Panel B of Table N.1 for the five main domestic appliances in terms of electricity use, and in Panel B of Table N.2 for other appliances. Standard errors are clustered by distribution utility using the wild cluster bootstrap-t (Cameron, Gelbach and Miller, 2008). The resulting confidence intervals are large given the small number of clusters, and typically include 0, so our results remain suggestive.

Point estimates are positive for our "age" index (older) when considering the main domestic appliances (it is 0 when considering all appliances in Table N.2) and for the age of each of the 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.[90] In fact, the supply side of the market 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 next sections, 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.

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 next sections, 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.

Finally, we showed in Table 4 that households reported adopting more energy-efficient light-

[89] 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 ten unidentified distribution 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), i.e., to a temporary effect if sales of electrical appliances decreased during the crisis, as we discuss in the next sections.

[90] 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.

bulbs 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 Southeast/Midwest and in the South between survey rounds. However, the difference-in-difference estimate suggests that adoption rates were in fact 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.

N.3 Utilization habits

The PPH surveys recorded utilization habits correlated with electricity use for many appliances. As above, we use the specification in equation (77) 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.

We display difference-in-difference estimates in Panel C of Table N.1 for the five main domestic appliances in terms of electricity use, and in Panel C of Table N.2 for other appliances. Standard errors are clustered by distribution utility using the wild cluster bootstrap-t (Cameron, Gelbach and Miller, 2008). The resulting confidence intervals are large given the small number of clusters, and typically include 0, so our results remain suggestive.

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 4. 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 alone could have generated enough savings to explain the long-term impact (22 kWh/month). [91] Finally, the only appliance for which we find a meaningful and statistically significant increase in utilization in Table N.2 is fans, the close substitute for air conditioners.

[91] 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).

Table N.1: Difference-in-difference results for appliances' quantity, characteristics, and utilization using PPH surveys (I)

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 High Power	Appliance Always Switched On Refrigerator	Appliance Always Switched On Freezer	Appliance Frequently Used Light	Appliance Frequently Used 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

The tables uses household-level data for eight distribution utilities in the Southeast/Midwest and two in the South from the two most recent rounds of the PPH surveys (1998-1999 and 2004-2005). It 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 estimating the specification in equation (77). 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 or Southeast/Midwest and 1999 or 2005). All regressions include distribution utility fixed effects, year fixed effects, and controls for income, income squared, number of household members, dwelling size, 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 of the variable on the household's level (class) of energy use and the remaining controls (separately in each round). Significance levels: *10%, **5%, ***1% (s.e. clustered by distribution utility using the wild-cluster bootstrap-t).

Table N.2: Difference-in-difference results for appliances' quantity, characteristics, and utilization using PPH surveys (II)

Panel A. Quantity

	Index (KKL)	Air Conditioner	Laundry	Iron	Dish Washer	Dryer	Microwave	Electric Oven	Fan	Heater
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$SE/MW \times Year2005$	**-.395**	**-.227**	**-.040**	**-.082***	**-.114**	**-.281**	**-.066**	**-.007**	**-1.077**	**-.031**
	(.542)	(.360)	(.057)	(.045)	(.167)	(.387)	(.102)	(.028)	(1.335)	(.053)
Average SE/MW 1999	-.005	.099	.532	.954	.049	.040	.212	.087	.811	.019
N	14,251	14,251	14,251	14,251	14,251	14,251	14,251	14,251	14,251	14,251

Panel B. Characteristics

	Index		Air conditioner	
	Age (KKL)	Type (KKL)	Age	Power (BTUs)
	(1)	(2)	(3)	(4)
$SE/MW \times Year2005$	**-.004**	**.082**	**-.234**	**1058.2**
	(.046)	(.119)	(.895)	(1259.3)
Average SE/MW 1999	-.018	.118	5.873	7823.043
N	14,206	14,206	888	805

Panel C. Utilization

	Index (KKL)	Appliance Frequently Used								
		Air Conditioner	Laundry	Iron	Dish Washer	Dryer	Microwave	Electric Oven	Fan	Heater
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$SE/MW \times Year2005$	**-.427**	**-.049**	**-.022**	**.008**	**-.014*****	**-.002**	**-.056**	**.010**	**.190*****	**.001**
	(.586)	(.036)	(.044)	(.064)	(.004)	(.007)	(.072)	(.008)	(.074)	(.007)
Average SE/MW 1999	.095	.040	.083	.130	.010	.004	.099	.014	.149	.002
N	14,251	14,251	14,251	14,251	14,251	14,251	14,251	14,251	14,251	14,251

The tables uses household-level data for eight distribution utilities in the Southeast/Midwest and two in the South from the two most recent rounds of the PPH surveys (1998-1999 and 2004-2005). It 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 estimating the specification in equation (77). 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. The *KKL indices* consider the average of the dependent variables shown in the columns of each panel, including the main appliances in Table N.1. When calculating the indices, each dependent variable is normalized by its 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 or Southeast/Midwest and 1999 or 2005). All regressions include distribution utility fixed effects, year fixed effects, and controls for income, income squared, number of household members, dwelling size, 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 of the variable on the household's level (class) of energy use and the remaining controls (separately in each round). Significance levels: *10%, **5%, ***1% (s.e. clustered by distribution utility using the wild-cluster bootstrap-t).

O Additional evidence on mechanisms of hysteresis: household expenditure surveys

We provide additional evidence on appliance holdings and vintage using self-reported household-level data from the last three rounds of the Brazilian Household Expenditure Surveys (POF, with rounds in 1996-1997, 2002-2003, and 2008-2009). The surveys, which contain household-level microdata (repeated cross-sections), were conducted by the Brazilian Geography & Statistics Institute (IBGE), which is also responsible for the National Census. All surveys were conducted between July of the base year and June of the following year. The 1996/1997 survey was based on a representative sample of the population of the largest metropolitan areas of Brazil. The two subsequent surveys were based on a representative sample of the whole Brazilian population. The microdata is not available at the municipality level and cannot be matched to the concession areas of distribution utilities. In our analysis, we restrict attention to the metropolitan areas of the seven states in the Southeast/Midwest and in the South, for which the samples are representative in all survey rounds.

The microdata contains data on the quantities of different types of appliances owned by the households and the year these appliances were bought. It does not have details about the model of these appliances, or whether the appliances were bought new or second-hand. Because these households who do not pay for electricity were not subject to the incentives of the energy-saving program, we exclude them from the main specifications. In POF surveys, a household may declare having more than one house. We discard second houses and restrict attention to the main housing unit. We truncate appliances' age at 20 years, because the year that an old appliance was bought is subject to severe measurement errors.

We use the different rounds of POF surveys to investigate any differential trend in appliances' quantity and age, using a difference-in-difference strategy as in Section 3 in the paper. Because we have three survey rounds, one just after the end of the energy saving program (2002-2003) and one more than six years later, we regress:

$$Y_{h,d,t} = \alpha_d + \sum_{t' \in \{2002,2008\}} \delta_{t'} \cdot \mathbb{1}(t = t' \,\&\, d \in \text{SE/MW}) + \gamma_t + X_{h,d,t} + \nu_{h,d,t} \tag{78}$$

where $Y_{h,d,t}$ is an outcome for household h from state d in survey round t. We control for state fixed effects α_d and a survey round fixed effect γ_t. The coefficients δ_{2002} and δ_{2008} are short-run and long-run difference-in-differences estimators under a common-trend assumption. We control for household characteristics, $X_{h,d,t}$, which may be correlated with different trends in appliance ownership. The vector of household characteristics includes income, income squared, number of household members and number of rooms. We also construct an appliance quantity index to avoid multiple-inference problems, as in the previous section.

Table O.1 presents the difference-in-differences estimates of the impact of the energy-saving program on the quantity (Panel A) and age (Panel B) of all electrical appliances recorded in POF surveys. Standard errors are clustered by distribution utility using the wild cluster bootstrap-t (Cameron, Gelbach and Miller, 2008). The resulting confidence intervals are large given the small number of clusters, so our results remain suggestive. We find close to zero long-run effects on the quantity of refrigerators and TVs. We find negative long-run effects on the quantity of freezers, which is consistent with the results based on the PPH surveys. We find small point estimates for a long-run effect on the age of appliances (older). We find a large increase in the average age of air conditioners (1.8 year older in the Southeast/Midwest in the long-run), which is substantially different from what we find using the PPH surveys, but these results should be interpreted with caution because of the small number of non-missing values for this variable in both surveys.

Table O.1: Difference-in-differences results for appliances' quantity and age using POF surveys

Panel A. Quantity

	Index (KKL)	Refrigerator	Freezer	TV	AC	Laundry	Iron	Dish Washer	Dryer	Microwave	Hair dryer	Sound System	Computer
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
$SE/MW \times Year2002$	**.019**	**.013**	**-.026**	**.044**	**.048**	**-.151**	**.025**	**.005**	**.055**	**.024**	**.028**	**.038**	**.019**
	(.016)	(.049)	(.027)	(.052)	(.031)	(.131)	(.120)	(.038)	(.080)	(.026)	(.214)	(.085)	(.029)
$SE/MW \times Year2008$	**-.019**	**.006**	**-.035**	**-.007**	**-.009**	**-.088**	**-.001**	**.020**	**.089**	**-.031**	**-.025**	**-.031**	**-.021**
	(.022)	(.011)	(.042)	(.084)	(.023)	(.097)	(.017)	(.055)	(.086)	(.020)	(.046)	(.027)	(.018)
Average SE/MW 1996	-.086	.982	.207	1.299	.132	.554	1.139	.077	.092	.196	.483	.755	.083
N	31113	31113	31113	31113	31113	31113	31113	31113	31113	31113	31113	31113	31113

Panel B. Age

	Index (KKL)	Refrigerator	Freezer	TV	AC	Laundry	Iron	Dish Washer	Dryer	Microwave	Hair dryer	Sound System	Computer
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
$SE/MW \times Year2002$	**.155****	**-.874**	**-.270**	**-.612***	**1.560****	**-.258**	**-.266**	**-14.999**	**.214**	**.469**	**.242**	**-.538*****	**-.057**
	(.069)	(.643)	(1.466)	(.332)	(.781)	(.251)	(.334)	(1994.2)	(1.139)	(.520)	(.487)	(.203)	(.343)
$SE/MW \times Year2008$	**.014****	**-.039**	**-.073**	**-.303**	**1.798**	**-.820**	**-.386**	**-1.712**	**-.143**	**.663**	**.137**	**.259**	**.227**
	(.006)	(.132)	(.905)	(.210)	(1.661)	(.567)	(.431)	(1365.9)	(4.372)	(.518)	(.641)	(.233)	(.318)
Average SE/MW 1996	-.026	8.024	5.336	4.378	6.46	7.038	5.119	6.032	6.215	3.257	5.76	4.687	1.784
N	31113	28908	5516	28198	2185	16018	27178	1460	2382	8093	10955	18544	7107

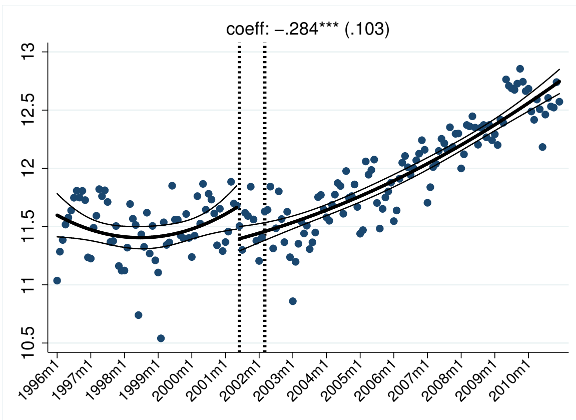
Household-level data for seven states (urban area only) in the Southeast/Midwest and in the South from the POF surveys in 1996-1997, 2002-2003 and 2008-2009. This table displays the difference-in-differences estimates of the impact of the energy-saving program on the quantity and age of all electrical appliances recorded in POF surveys, using the specification in equation (78). Each column corresponds to a regression for a different appliance. *Panel A* displays the results for the quantity of appliances owned by household, and *Panel B* displays the results for the age of appliances owned by households. The *Indices (KKL)* shown in the first columns are the average of the appliances shown in the columns. When calculating the indices, each dependent variable is normalized by the average and standard deviation of each variable in the South in 1996 (Kling, Liebman and Katz, 2007). To obtain these indices, we input missing values with the mean of the cell group to which the household belongs (South or Southeast/Midwest and 1996 or 2002 or 2008). We include only households who pay for electricity (non-theft). All regressions contain distribution utility fixed effects, year fixed effects, and controls for income, income squared, number of household members, and number of rooms. We input missing values in income using a linear regression of income on the remaining controls (separately for each survey round). Significance levels: *10%, **5%, ***1% (s.e. clustered by state using the wild-cluster bootstrap-t).

P Additional evidence on mechanisms of hysteresis: national sales of appliances

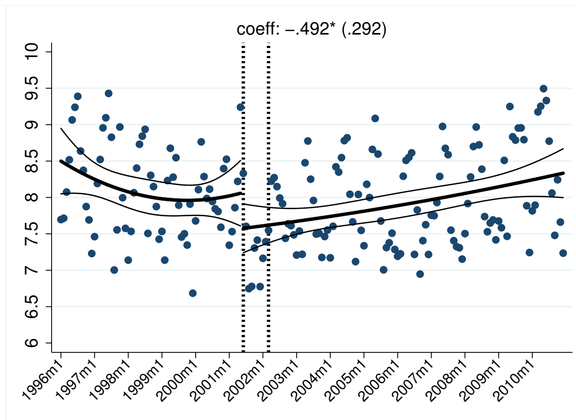
The figures in this section display the time-series in national sales of several domestic appliances in Brazil. The data are estimates that we obtained from Whirlpool, a leading manufacturer, which produces those estimates for its own market strategy. The manufacturer did not share with us the estimation methodology it used. In each figure, we plot the raw data (in logs) and a quadratic fit on each side of the start of the crisis (June 2001), and we display the estimated change in sales at the time of the crisis from a regression discontinuity design using those quadratic fits. The Southeast/Midwest is by far the largest market for domestic appliances in Brazil (more than 50%). We find no evidence of an increase in sales for any of the domestic appliances. In contrast, we find evidence of a discontinuous decrease in national sales for several of them, in particular washing machines, freezers, microwaves, and dishwashers.

Figure P.1: Log sales of different domestic appliances I (all Brazil, Whirlpool estimates)

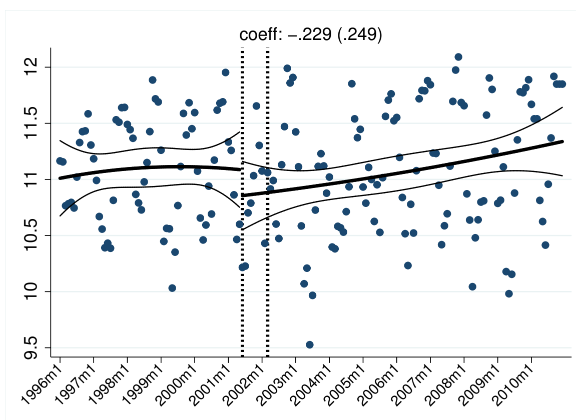
(a) Washing machine



(b) Dryer



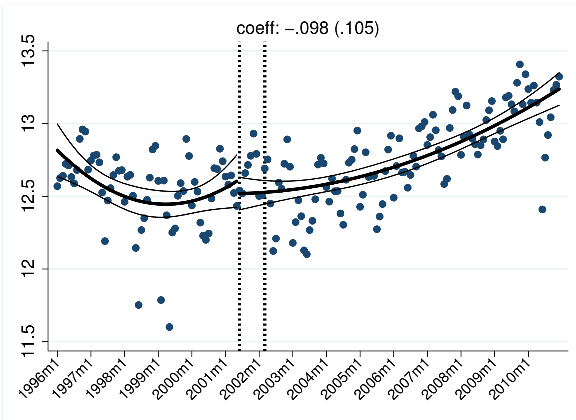
(c) Air conditioning



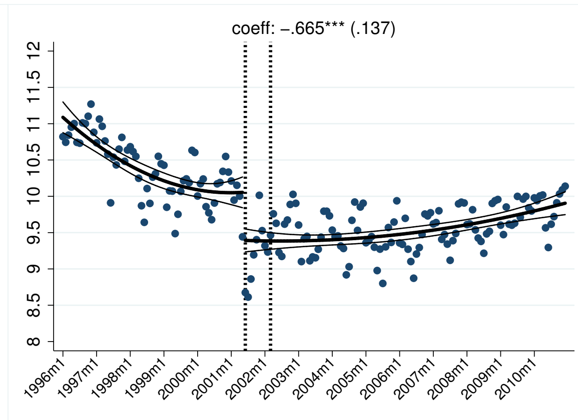
The panels display the time series in national sales of several domestic appliances in Brazil. The data are estimates that we obtained from Whirlpool, a leading manufacturer, which produces those estimates for its own market strategy. The manufacturer did not share with us the estimation methodology it used. In each figure, we plot the raw data (in logs) and a quadratic fit on each side of the start of the crisis (June 2001), and we display the estimated change in sales at the time of the crisis from a regression discontinuity design using those quadratic fits. The Southeast/Midwest is by far the largest market for domestic appliances in Brazil (more than 50%).

Figure P.2: Log sales of different domestic appliances II (all Brazil, Whirlpool estimates)

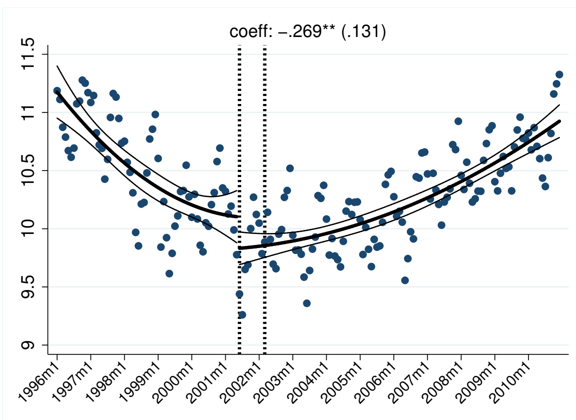
(a) Refrigerator



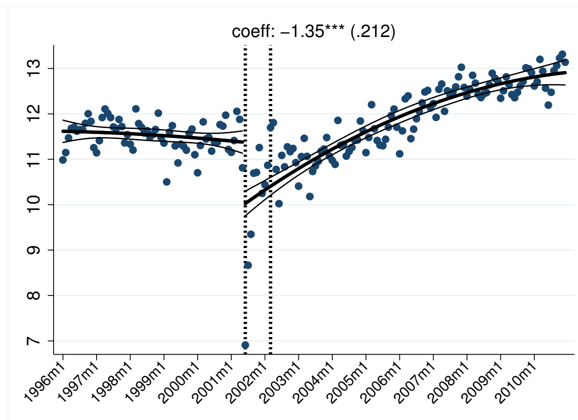
(b) Freezer (vertical)



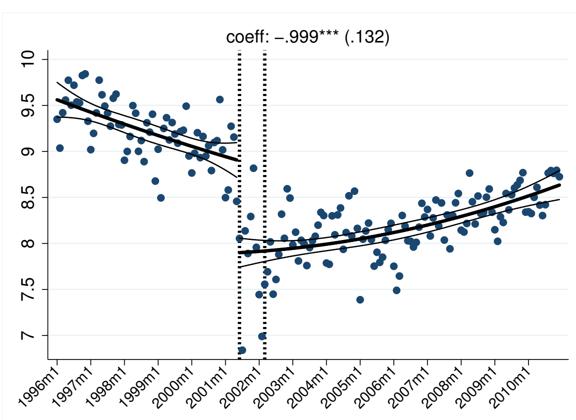
(c) Freezer (horizontal)



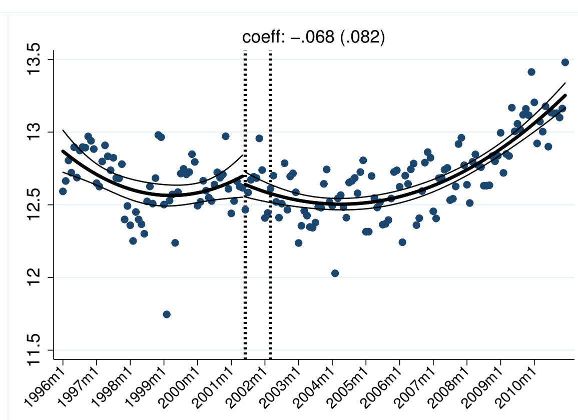
(d) Microwaves



(e) Dishwasher



(f) Gas stove (placebo)



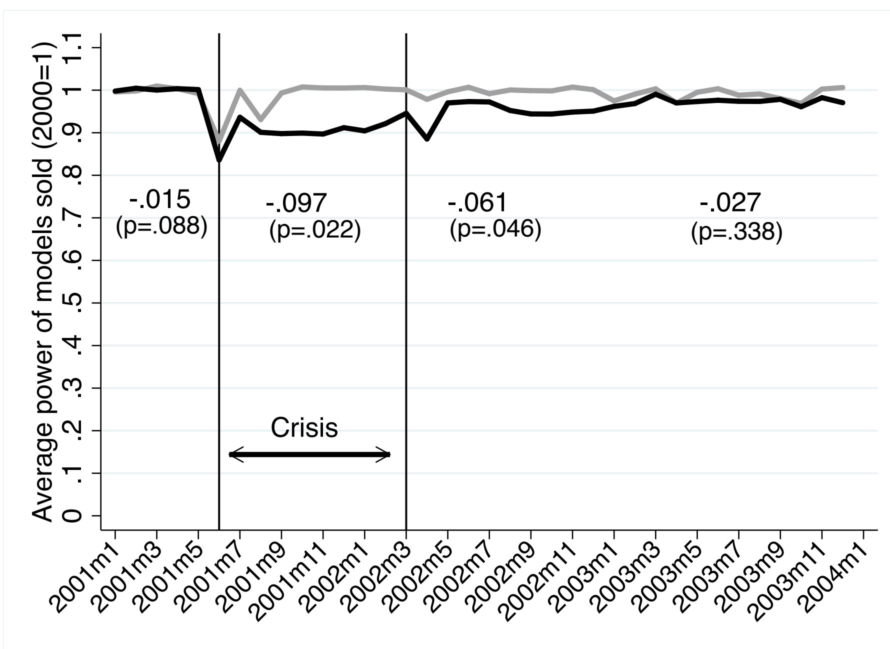
The panels display the time series in national sales of several domestic appliances in Brazil. The data are estimates that we obtained from Whirlpool, a leading manufacturer, which produces those estimates for its own market strategy. The manufacturer did not share with us the estimation methodology it used. In each figure, we plot the raw data (in logs) and a quadratic fit on each side of the start of the crisis (June 2001), and we display the estimated change in sales at the time of the crisis from a regression discontinuity design using those quadratic fits. The Southeast/Midwest is by far the largest market for domestic appliances in Brazil (more than 50%). Gas stoves constitute a placebo because they do not use electricity.

Q Additional evidence on mechanisms of hysteresis: sales of electric showers

Finally, Figure Q.1 uses data on the monthly sales of all the models of electric showers sold by Fame, a leading manufacturer, in each Brazilian state between January 2000 and December 2003. The data include the power (wattage) of each model, which is the only relevant measure of electric showers' propensity to use electricity. Figure Q.1 displays the average power of electric showers sold in each month in the Southeast/Midwest and in the South, normalized to the same months in 2000. It also displays difference-in-differences estimates in each time period (early 2001, crisis, rest of 2002, 2003) from regressing the logarithm of average power on dummies for each state, dummies for each time period, and those time-period dummies interacted with an indicator for weather distribution utilities in the state were subject to the energy-saving program during the crisis. Standard errors are clustered by state using the wild cluster bootstrap-t (10 clusters).

The average power of electric showers sold decreased by about 10% during the crisis in the Southeast/Midwest compared to the South. We do not find any evidence of persistence. Note that, in the Southeast/Midwest, Fame also started to have relatively fewer sales compared to the South (not shown; coefficient estimates are not significantly different from 0 during and after the crisis).

Figure Q.1: Average power of electric showers sold (one manufacturer, Fame)



The figure displays the average power of electric showers sold by Fame in each month in the Southeast/Midwest and in the South, normalized to the same months in 2000. It also displays difference-in-differences estimates in each time period (early 2001, crisis, rest of 2002, 2003) from regressing the logarithm of average power on dummies for each state, dummies for each time period, and those time-period dummies interacted with an indicator for whether distribution utilities in the state were subject to the energy-saving program during the crisis. P-values are reported based on standard errors clustered by state using the wild cluster bootstrap-t (10 clusters).