Pigouvian Taxes with Intermittent Billing and Habits☆

DRAFT. Preliminary and Incomplete.

Ben Gilbert[a,*], Joshua S. Graff Zivin[b]

[a] *University of Wyoming, 1000 E University Ave, Laramie WY, 82071, United States*
[b] *University of California, San Diego and NBER, 9500 Gilman Dr, La Jolla CA, 92093-0519*

Abstract

This paper derives an optimal Pigouvian tax framework for goods whose consumption is habit-forming and socially costly, and whose true costs are only intermittently salient. Habit formation is an empirical feature of several prominent goods that have external costs, such as gasoline, electricity, and unhealthy foods. Habits are exacerbated when payment for consumption of these goods occurs intermittently and therefore prices are usually not salient; electricity use is typically billed monthly, and driving and food consumption decisions are often made at higher frequency than payment - particularly if consumers pay with credit cards that are billed monthly. The intermittency of payment creates time periods in which prices of individual decisions are more and less salient. Optimization errors when prices are not salient lead to inefficiently large habit stocks. When bills arrive, households minimize the consequences of future errors by making new habits in the present. An optimal Pigouvian tax must match the timing and degree of price inattention and is therefore not practical to implement. We characterize a second-best constant tax and the deadweight loss associated with it. Using a household-level panel of daily electricity use and monthly bills, we provide evidence that residential electricity consumption is consistent with habit formation and we show how the habit parameters change when prices are more and less salient.

Keywords: Electricity Demand; Salience; Habit Formation; Inattention; Smart Meters

☆We thank numerous seminar participants for helpful comments. We are grateful to San Diego Gas & Electric for providing us access to their data. Any errors are our own.

*Corresponding author. ph: 1 307 766 6315, fax: 1 307 766 5090

Email addresses: bgilbe10@uwyo.edu (Ben Gilbert), jgraffzivin@ucsd.edu (Joshua S. Graff Zivin)

1. Introduction

Pigouvian taxes are designed to align marginal benefits with marginal social costs. Consumers who make optimization errors when prices are not fully salient, however, may not use their true marginal benefit as part of their decision criteria. Furthermore if agents form habits over the consumption of certain goods then marginal benefits are not independent of past and (expected) future consumption. In this case optimization errors that occur periodically will cause intertemporal externalities. Both of these behaviors - habit formation and price inattention - are relevant features of several classes of goods that also have external costs associated with their consumption. Previous literature suggests these features may also be prevalent in gasoline and unhealthy food consumption, which impose external pollution damage and health costs on society. In this paper we show empirically that these features are prevalent in daily household electricity consumption.

Further, we derive a Pigouvian tax framework for market settings that exhibit these features. A key issue is that markets for these types of goods often fail to temporally align the actions of payment and consumption. This alignment may be necessary to improve price attention given cognitive limitations. Limited empirical studies on situations with temporally-aligned price visibility and consumption decisions have shown that consumption is significantly lower if true costs are visible at the moment the decision is made. Examples include pay-as-you-go electricity systems in the United Kingdom, an increasing number of real-time electricity information trials in U.S. households, and automobile dashboard displays of fuel consumption rates.

How will corrective taxes fare in this context? This paper derives a framework for Pigouvian taxes in an environment of intermittent price salience and habit formation. We show that the optimal Pigouvian tax varies over time with the level of price salience, and is therefore unlikely to be feasible. We provide formulas for the second-best constant tax and the associated deadweight loss from being second-best. These formulas are based on parameters of salience, habit formation, price elasticity, and time preference that are either estimable or available in the literature.

Lastly, we provide estimates of the habit formation parameters for the case of household electricity consumption, on which there are not preexisting empirical estimates to our knowledge. We use a data set of hourly smart meter readings from 10,000 households in the suburbs of San Diego, CA, to calculate daily use and to construct additional household covariates that proxy for preferences, behavior, and durable goods stock. Consistent with our model, our results suggest a strong persistence in the force of habit - in other words, a habits-as-durables model over a short memory habit model - implying potentially large long-run impacts that take time to take effect following permanent price changes, but small effects from transitory price changes. Our evidence also suggests that information alone can change habits if it is sustained; in the days after households receive a new monthly bill their accumulation of habit stock accelerates - they are able to make new habits more quickly - but the acceleration disappears as the billing cycle continues and the impact of the information in the bill fades.

2. Literature

Gaps in the literature occur at the intersection between dynamic implications of hysteretic consumption (e.g., rational habits) on the one hand, and cognitive and information problems like price salience, inattention, and awareness on the other hand. As we discuss below, the price salience literature has a dearth of work on dynamic consequences, and the habit formation problem has not been thoroughly analyzed in the context of cognitive/behavioral limitations. Both of these have important implications for electricity markets and conservation. Two exceptions are Allcott, Mullainathan, and Taubinsky (2014) and Taubinsky (2013). Taubinsky (2013) proposes a model in which habits arise endogenously as the result of inattention; in our model these two features are exogenous and separate but coexist within the same agent. Dynamic consequences of inefficient durable goods stock accumulation from inattention to future prices has been studied by Allcott, Mullainathan, and Taubinsky (2014), who show that Pigouvian taxes can have both external and internal corrective benefits in this case.

2.1. Price Salience and Information

Consumer inattention in contexts of limited price salience has been the subject of a swiftly growing literature in economics. Overconsumption because of opaque prices has been demonstrated in the contexts of beer, groceries, bank overdrafts, cell phone overages, automatic bill payments, lumpy future costs like car and home repairs (Sexton, 2012; Grubb, 2012; Stango and Zinman, 2011; Karlan, et., al., 2010; Finkelstein, 2007; Chetty, et., al., 2007). Della Vigna (2009) describes attention as a scarce resource; when this resource is not fully allocated to cognitively processing a complicated or unclear price, the price seems lower than it really is to the consumer. There is even recent evidence that the availability of other (economic) resources can influence how the attention resource is allocated (Shah, et., al., 2012).

A large and related literature has investigated the potential for energy conservation to achieve pollution abatement goals. Assessments of this potential vary widely (McKinsey, 2007; Jaffe & Stavins, 1994) and more recently studies have evaluated the prospect of behavioral "nudges", such as goal-setting, social norms, and personalized feedback, to change energy use patterns (Hsiaw and Harding, 2013; Allcott, 2011; Allcott & Mullainathan, 2010). Consumption reductions of between 2% and 5% have been documented.

The lack of salient price and quantity information has also been discussed as a cause of inefficient electricity use (Jessoe and Rapson, 2014; Gilbert and Graff Zivin, 2013). Jessoe & Rapson (2014), for example, show that households whose consumption information is updated in real-time and displayed prominently by in-home devices are much more responsive to price signals than control households. Gilbert and Graff Zivin (2013) find that households reduce electricity consumption by about 1% for several days after receiving a routine monthly bill. Allcott & Rogers (2012) find that household electricity consumption is also responsive to personalized reminder messages but that the effects of these interventions can erode over time.

2.2. Habits in Consumption

The causes of persistence in the time series of consumption at the macroeconomic level have been debated at least since Hall (1978), who showed that the permanent income hypothesis implies an AR(1) process in aggregate consumption. Another explanation has been attributed to habits, which imply further lags of consumption the stronger is the force of habit. Habits in a myopic setting were addressed by Pollak (1970), and a model of rational habit formation with forward-looking agents was introduced by Spinnewyn (1981). A key feature of habits is that they make utility time-inseparable, as the marginal utility of today's consumption is affected by past consumption (or a stock of habits accumulated from past consumption).

Solving explicitly for demand functions has proven unwieldy because of the forward- and backward-recursions introduced by time-inseparability and forward-looking agents. Demand becomes a function of all past consumption (or state variables capturing this series) as well as expectations of all future realizations of prices, incomes, and other demand shifters. This is demonstrated in Spinnewyn (1981) and Browning (1991). As noted by Scott (2012), this feature of solved demands makes empirical work difficult. The length of these recursions, however, depends on the duration of the habits. If habits are captured by only one lag of consumption (short memory), then demand depends on one lag of consumption and the entire future series of expected prices. If habits are "durable", so that all past consumption describes a single habit stock or state variable at any given time, then the demand function depends on all lags of consumption and only one future lead of price. For an excellent discussion of this literature and the empirical implications, see Scott (2012).

A feasible approach to dealing with functions of an infinite series of variables suggested by Browning (1991), Becker, et., al., (1994) and Scott (2012) is to truncate the lags or leads. We follow this approach in our empirical section. Using this approach, Scott (2012) finds evidence of habit formation and sensitivity to expected future price changes in gasoline demand, concluding that carbon taxes on gasoline consumption may be more effective in the long run than has been previously assumed using standard demand elasticity estimates. Weak evidence of habit persistence in household electricity use has been found by Leth-Petersen (2007), Heien and Durham (1991), and Sexauer (1977), all of which use either monthly bills or consumer expenditure surveys. Our study is the first to find strong evidence at the daily level, which is the relevant frequency to study the relationship between habit formation and time-varying price salience.

3. Model

Utility in each period depends on consumption of a clean numeraire good y_t, a dirty habit-forming good x_t, and a reference or habit stock of past consumption of the dirty good z_t. The habit stock is a moving average of past consumption as in

$$z_t = \rho x_{t-1} + (1-\rho) z_{t-1} \tag{1}$$

Consumption of the dirty good produces social damages that are proportional to the level of consumption and last for only the current period:

$$D_t = \phi x_t. \tag{2}$$

For simplicity we analyze the case in which period utility is quasilinear:

$$U_t = u_t(x_t - \alpha z_t) + y_t \tag{3}$$

where subscripts on U_t and u_t are only used to keep track of the time period (the functional forms are assumed to be the same in all periods). In this formulation, consumption of x_t across time periods are "adjacent complements" in the sense that as z_t increases from past consumption, the marginal utility of x_t rises in period t and consumption of x_t increases *ceteris paribus*. This provides the agent a lever with which to mitigate potential future overconsumption by reducing consumption in the present in order to impose better habits on the future[1]. We ignore saving and borrowing so that the household faces a budget constraint in each period:

$$m = px_t + y_t \tag{4}$$

This simplification allows us to focus on the dynamic relationship from habit formation rather than savings and borrowing. It also allows us to avoid cases where the household would be so concerned about future overconsumption that it would "raid" its savings in the present and leave nothing to be wasted in by suboptimal decisions in the future. We do not think this case is empirically relevant for household energy consumption.

3.1. The Social Optimum

A three-period analysis is sufficient to gain insight from the model. Assuming a rate of time preference equal to the interest rate, the social planner chooses a path $\{x_t, y_t\}_{t=1}^3$ to maximize the intertemporal social welfare function

$$W = \sum_{t=1}^{3} \beta^{t-1} \left[U_t(x_t, y_t, z_t) - \phi x_t - \lambda_t(px_t + y_t - m)\right] \tag{5}$$

subject to equation (1) and z_1 given. The optimal path is defined by the first order conditions:

$$\begin{gathered} u_1' - \alpha\rho\beta u_2' - \alpha\rho(1-\rho)\beta^2 u_3' = p + \phi \\ u_2' - \alpha\rho\beta u_3' = p + \phi \\ u_3' = p + \phi \end{gathered} \tag{6}$$

One implication of (6) is that absent problems of salience and inattention, a standard Pigouvian tax of $\tau = \phi$ would induce households to consume at the social optimum.

3.2. The Private Optimum

Consistent with empirical evidence in the literature, we assume that during periods in which the price of x is not salient, households overconsume and are less responsive to prices than when they are fully salient. We represent this by having households undervalue the

[1]Habits are not required to introduce a dynamic element to the problem; introducing borrowing and savings or a single permanent income constraint will provide a similar lever. We focus on habits because they are an empirical feature of the polluting goods we are interested in, and because intermittent price salience has important implications for habit formation.

true price by a salience factor $\theta \in (0,1)$ and optimize as if their period budget constraint was:

$$m = \theta p x_t + y_t \tag{7}$$

We will assume that prices are salient in period 1, but not in periods 2 or 3. The household's period 2 problem becomes:

$$\begin{gathered} max_{x_2,y_2,x_3,y_3,z_3} \sum_{t=2}^{3} \beta^{t-2} \left[U_t(x_t, y_t, z_t) - \lambda_t(\theta p x_t + y_t - m)\right] \\ s.t.\ z_3 = \rho x_2 + (1-\rho) z_2 \\ z_2 \text{ given} \end{gathered} \tag{8}$$

The consumption path chosen subject to optimization errors induced by θ is defined by the first order conditions:

$$\begin{gathered} u_2' - \alpha\rho\beta u_3' = \theta p \\ u_3' = \theta p \end{gathered} \tag{9}$$

Note that if households maintained the same θ in all periods, in principle the optimal consumption path defined in (6) could be induced by an extra large tax $\tau_2(\theta) = \tau_3(\theta) = \frac{1-\theta}{\theta} p + \frac{\phi}{\theta} > \phi$.

Let the consumption path defined by (9) be

$$\begin{aligned} \tilde{x}_2 &= \tilde{x}_2(z_2, p; \theta) \\ \tilde{x}_3 &= \tilde{x}_3(z_2, p; \theta) \\ \tilde{y}_2 &= \tilde{y}_2(z_2, p; \theta) \\ \tilde{y}_3 &= \tilde{y}_3(z_2, p; \theta) \end{aligned} \tag{10}$$

Notice that $\tilde{x}_t$ is larger than the privately optimal demand without salience effects, whereas $\tilde{y}_t$ is smaller.

In period 1 the household receives a price signal. We assume the household is sophisticated enough to know its demands in periods 2 and 3 are those given in (10). When the household chooses x_1, it will take account of how that choice will influence these future demands through z_2. The houshold's period 1 problem is then

$$max_{x_1,y_1,z_2} U_1(x_1, y_1, z_1) - \lambda_1(p x_1 + y_1 - m) + \sum_{t=2}^{3} \beta^{t-1} \left[U_t(\tilde{x}_t, \tilde{y}_t, z_t)\right] \tag{11}$$

subject to equation 1 and z_1 given. The first order conditions for this problem can be reduced to

$$\left[u_1' - \alpha\rho\beta u_2' - \alpha\rho(1-\rho)\beta^2 u_3'\right] + \rho\beta\frac{\partial\tilde{x}_2}{\partial z_2}\left[u_2' - \alpha\rho\beta u_3' - p\right] + \rho\beta^2\frac{\partial\tilde{x}_3}{\partial z_2}\left[u_3' - p\right] = p \qquad (12)$$

The first term in brackets is the marginal utility across time taking habits into account and would be equal to the price if salience effects did not exist. The second two terms are negative, and measure the extent of the deviation from a private optimum in periods 2 and 3 due to salience effects. Denote the period 1 demand that solves equation (12) as $\tilde{x}_1(z_1, p; \theta)$. This is lower than the privately optimal demand would be in the absence of salience effects.

By substituting the conditions in (9) into (12) and simplifying, we obtain

$$u_1' - \alpha\rho\beta u_2' - \alpha\rho(1-\rho)\beta^2 u_3' = p\,(1+\delta) \qquad (13)$$

where $\delta = \rho\beta(1-\theta)\left(\frac{\partial\tilde{x}_2}{\partial z_2} + \beta\frac{\partial\tilde{x}_3}{\partial z_2}\right) = \rho\beta(1-\theta)\,(\alpha + \alpha(1-\rho)\beta)$. With quasilinear utility and linear demand functions, terms $\frac{\partial\tilde{x}_t}{\partial z_2}$ are constant demand shifters, so δ is a constant.

3.3. Pigouvian Taxation

By comparing the first order conditions in (9) and (13) to the conditions at the optimum in (6), we can derive a set of optimal taxes that will direct the household to consume optimally in every period. These are stated as a proposition:

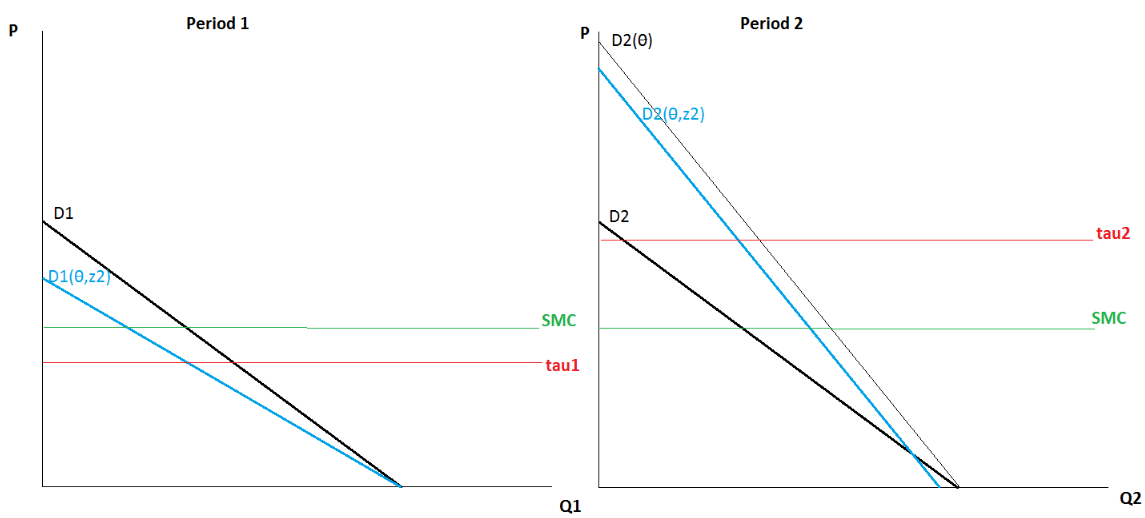
Proposition 1. *The optimal Pigouvian tax is time-varying, with*

$$\begin{aligned} \tau_2(\theta) = \tau_3(\theta) &= \frac{1-\theta}{\theta}p + \frac{\phi}{\theta} > \phi \\ \tau_1(\theta) &= \frac{\phi}{1+\delta} - \frac{\delta}{1+\delta}p < \phi \end{aligned} \qquad (14)$$

Notice that $\tau_1(\theta)$ is less than the marginal social damage and can even be a subsidy. The household is already underconsuming relative to the optimum in period 1 in order to preempt overconsumption in future periods, so taxes can be lower than marginal external costs. It is also important to note the only behavioral parameter affecting $\tau_2(\theta)$ and $\tau_3(\theta)$ is the inattention parameter θ, while the habit parameters ρ and $\frac{\partial\tilde{x}_t}{\partial z_2}$ (or α) both affect $\tau_1(\theta)$ through δ. This is because in period 1, the household is using it's habit formation process to manipulate future consumption and the regulator must optimally address that manipulation process. In later periods, however, the household makes a boundedly rational choice about habit formation; the regulator simply needs to correct the "boundedness" of that rational choice in addition to the social damage. So the later taxes need not directly address the habit process except through adjustments for the lack of salience.

These taxes are displayed in Figure 1. The lines labeled $D1$ and $D2$ are the privately optimal demand functions if no salience effects exist. Because prices are insalient in period 2, the demand curve rotates up to $D2(\theta)$ - more is consumed at every price, and demand is less responsive to prices. However, in anticipation of this the household consumes less in period 1. This is captured by the downward rotation to demand curve $D1(\theta, z_2)$ in period 1. This reduction in consumption also reduces the habit stock inherited in period 2, z_2, which shifts the period 2 demand curve inward to $D2(\theta, z_2)$. These final demands in blue are used by the household in deciding the consumption bundle. The optimal taxes defined in Proposition 1 lead the household to consume where there true marginal benefits given by $D1$ and $D2$ intersect the social marginal cost.

Figure 1: Optimal Time-varying Pigouvian Taxes



This set of taxes is not feasible to implement in practice however because the regulator needs to know the degree of time-varying inattention to true prices at all time periods and have the administrative capacity to implement such a variable tax. With these limitations in mind we derive a second-best constant tax, discuss its properties, and show how much surplus is given up by moving from the optimal time-varying tax. The size of that loss can help inform regulators about how much to invest in systems to improve price salience.

Let $\tilde{W}(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}, \mathbf{z}) = \tilde{W}(z_1, p; \theta)$ be the social welfare function from (5) evaluated at the consumption path given in (10) and the solution to (12) (e.g., the privately optimal time path with periodic salience and habit formation). Suppose a small tax τ is added to a previously untaxed environment so that we evaluate $\tilde{W}(z_1, p+\tau;\theta)$. The second-best optimal constant tax solves $\frac{\partial \tilde{W}}{\partial \tau} = 0$. The slope of the demand curves are constant in our case, so let's define

$$\begin{aligned} \tfrac{\partial \tilde{x}_1}{\partial \tau} &= a \\ \tfrac{\partial \tilde{x}_2}{\partial \tau} &= b \\ \tfrac{\partial \tilde{x}_3}{\partial \tau} &= c \end{aligned} \tag{15}$$

Proposition 2. *The second-best constant tax rate is given by*

$$\tau^{\star} = \frac{\frac{a}{\theta}(1+\delta)}{\frac{a}{\theta}(1+\delta)+\beta(b+\beta c)}\tau_1(\theta) + \frac{\beta(b+\beta c)}{\frac{a}{\theta}(1+\delta)+\beta(b+\beta c)}\tau_2(\theta) \tag{16}$$

where $\tau_1(\theta)$ and $\tau_2(\theta)$ were defined inProposition 1.

Proof. By plugging the consumption path given in (10) and the solution to (12) into the social welfare function and differentiating with respect to τ, we obtain

$$\frac{\partial \tilde{W}}{\partial \tau} = 0 = a\left[u_1' - \alpha\rho\beta u_2' - \alpha\rho(1-\rho)\beta^2 u_3' - p - \phi\right] + \beta b\left[u_2' - \alpha\rho\beta u_3' - p - \phi\right] + \beta^2 c\left[u_3' - p - \phi\right]$$

From the first order conditions in (9) and (13) (modified to include the introduction of a small tax) we can simplify this to

$$0 = a\left[(p+\tau)(1+\delta) - p - \phi\right] + \beta b\left[\theta(p+\tau) - p - \phi\right] + \beta^2 c\left[\theta(p+\tau) - p - \phi\right]$$

Solving for τ and plugging in the formulas for $\tau_1(\theta)$ and $\tau_2(\theta)$ gives the result. ■

Proposition 2 says that the best we can do with a constant tax is a weighted average of the optimal variable taxes, where the weights depend on parameters of the habit process, the slopes of the demand curves, the degree of price salience, and the discount rate. Notice that as θ approaches 1, $\tau^{\star}$ approaches ϕ because τ_1 and τ_2 also approach ϕ.

The size of this tax depends crucially on the behavioral parameters of the habit process. If new habits are easier to form (if ρ is large), and/or if future demand is more responsive to that new habit (if the $\frac{\partial \tilde{x}_t}{\partial z_2}$ are large), then the household makes larger corrections in period 1 on its own. In this case, the constant tax must be balanced more towards correcting under-consumption in period 1. Larger values of these parameters cause δ to also be large. With a large δ, τ_1 becomes more heavily weighted and τ_1 itself becomes smaller, so the constant tax $\tau^{\star}$ also becomes smaller. This does not necessarily mean that stronger habit persistence is better for welfare - in fact we will show the opposite is true in the next section. Intuitively, with larger habit persistence there is more imbalance in consumption across price salient and insalient time periods. Part of the pollution control and overconsumption purpose of the tax is sacrificed to prevent the household from consuming too little when prices are salient. A large δ will also occur if the discount rate is low (so that the household adjusts its habits out of concern for the future and needs less of a tax incentive). A large ρ indicates that new habits are easier to form because the habit stock is more responsive to more recent consumption. A large $\frac{\partial \tilde{x}_t}{\partial z_2}$ indicates that future consumption is more responsive to changes in the habit stock, so the household can have a larger impact on future (suboptimal) consumption by changing its habits in the present. Likewise if δ is small because habits are stagnant (low ρ) or consumption is not very responsive to the habit stock (low $\frac{\partial \tilde{x}_t}{\partial z_2}$), the emphasis in the constant tax can be shifted back towards reducing pollution and overconsumption in later periods, and the tax is larger.

3.4. Efficiency Cost

We now turn to deriving the efficiency cost of the second-best constant tax. Although $\tau^\star$ is chosen to minimize the deadweight loss from over- and under-consumption by definition, the remaining losses have policy relevance. The size of the deadweight loss (and of $\tau^\star$) depends on behavioral parameters for both salience and habit formation. To the extent that alternative policy tools (such as information technology investments) can affect these parameters, they can influence the size of the deadweight loss. For example, smart electricity grid technologies that provide real-time price information to households require large fixed costs, but the new information that households receive may alter their inattention and habit formation behavior. Therefore an estimable expression for this loss and how it depends on these parameters may be important for these investment decisions.

We will use the concept of equivalent variation to derive an expression for the excess burden on consumers when the regulator imposes a tax that cannot correct all market failures. The loss will be calculated as the net present value of the wealth society would be willing to forgoe in order to avoid the imperfect tax instrument (relative to the optimum), net of any changes in tax revenue between the two policies. Because the reference (optimal) tax level is different in different time periods, we will represent society's expenditure function and indirect welfare function as depending on a vector sequence of tax-inclusive prices (even though the tax-inclusive price will be constant for a constant tax). The expression for excess burden is

$$EB(\underline{\tau}^{\star}) = m(1+\beta+\beta^2) \\ - e\left(p+\tau_1(\theta), p+\tau_2(\theta), p+\tau_3(\theta), \tilde{W}(p+\tau_1^{\star}, p+\tau_2^{\star}, p+\tau_3^{\star}, m, z_1; \theta)\right) \\ - \left(R(\underline{\tau}^{\star}) - R(\underline{\tau}(\theta))\right) \quad (17)$$

where $\tau_1^{\star} = \tau_2^{\star} = \tau_3^{\star} = \tau^{\star}$ and $\underline{\tau}$ is a vector of taxes. The function $e(\cdot)$ is society's expenditure function. It is the amount of wealth at the optimal set of taxes that would be required to achieve the level of welfare obtained under the second-best constant tax. The difference between $e(\cdot)$ and the net present value of income is the amount of wealth society would be willing to forgoe to retain the optimal tax structure. The last term in parantheses measures the tax revenue difference between the two policies. For example,

$$R(\underline{\tau}^{\star}) = \tau_1^{\star}\tilde{x}_1(p+\tau_1^{\star}, z_1) + \beta\tau_2^{\star}\tilde{x}_2(p+\tau_2^{\star}, z_2(x_1)) + \beta^2\tau_3^{\star}\tilde{x}_3(p+\tau_3^{\star}, z_2(x_1)).$$

Following Auerbach (1985), we derive a more convenient expression for $EB(\underline{\tau}^{\star})$ using a second-order Taylor expansion. We calculate the expansion around the optimal tax vector $\underline{\tau}(\theta)$.

$$EB(\underline{\tau}^{\star}) \approx \frac{\partial EB}{\partial \underline{\tau}'}(\underline{\tau}(\theta)) \cdot (\underline{\tau}^{\star} - \underline{\tau}(\theta)) + \frac{1}{2}(\underline{\tau}^{\star} - \underline{\tau}(\theta))' \cdot \frac{\partial^2 EB}{\partial \underline{\tau}\partial \underline{\tau}'}(\underline{\tau}(\theta)) \cdot (\underline{\tau}^{\star} - \underline{\tau}(\theta)) \quad (18)$$

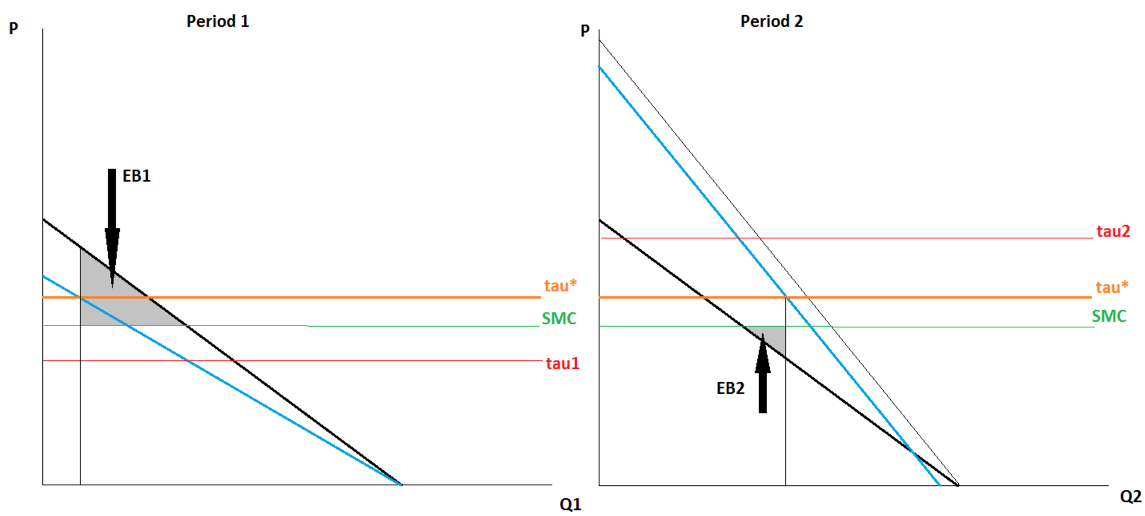
With quasilinear utility and linear demand curves, $EB(\underline{\tau}^{\star})$ will simply be the net present value of a sequence of deadweight loss triangles. Because of the habit process, however, deviations from the optimal tax sequence in a given period may cause the demand curves in other periods to rotate as they adjust to the new habit stocks. So the heights of these deadweight loss triangles need to adjust for the rotations that occur because of the habit stock.

The envelope theorem guarantees that the first-order terms in the Taylor expansion will be zero when evaluated at the optimal tax sequence $\underline{\tau}(\theta)$. The second order terms require taking derivatives of marginal utilities in each period with respect to tax changes in the same and all other periods. In our context, a change in the marginal utility in the current period because of a tax change in a later period results in a rotation in the current period demand curve, or a change in the willingness to pay for present consumption. After extensive algebra, we can show that

$$EB(\underline{\tau}^\star) \approx \frac{1}{2}\left(\underline{\tau}^\star - \underline{\tau}(\theta)\right)' \begin{bmatrix} \frac{\partial x_1}{\partial \tau_1} & -\alpha\rho\beta\frac{\partial x_1}{\partial \tau_1} & -(\alpha\rho\beta)^2\frac{\partial x_1}{\partial \tau_1} \\ 0 & \beta\frac{\partial x_2}{\partial \tau_2} & 0 \\ 0 & 0 & \beta^2\frac{\partial x_3}{\partial \tau_3} \end{bmatrix} \left(\underline{\tau}^\star - \underline{\tau}(\theta)\right) \tag{19}$$

The deadweight loss described in (19) is pictured in Figure 2 (for the first two periods).

Figure 2: Excess Burden of the Second-best Constant Tax



As drawn, the excess burden in period 1 is larger than in period 2, but the period 2 losses are repeated for as many periods as the prices remain insalient. The total excess burden in (19) is increasing in the habit formation parameters ρ and α, holding taxes constant. When the taxes in periods 2 and 3 are lowered from their optimal values to meet the second-best constant rate, overconsumption becomes a problem again and the household makes larger adjustments in period 1, rotating the period 1 demand curve further downward. Households with larger habit parameters will make larger period 1 adjustments. The more the demand curve rotates downward, the greater the underconsumption problem in period 1, and the greater the period 1 deadweight loss, which increases total losses.

A useful feature of (19) is that all the parameters in the expression are in principle estimable or recoverable from recent studies in the literature. For the residential electricity consumption example, the slopes of salient and insalient demand curves can be inferred from results by Jessoe and Rapson (2014), salience parameters can be inferred from Gilbert and Graff Zivin (2013) or Allcott & Rogers (2012), marginal social damages of various pollutants are available from, among others, Graff Zivin, Kotchen, and Mansur (2012). We provide estimates of the habit formation parameters in the remainder of this paper.

4. Habits in Daily Electricity Use

Our model has significant implications for the electricity industry and the multi-billion-dollar "smart grid" investments that are being implemented by many utilities with the goal of gaining control of demand. In the status quo electricity consumption environment, information is limited, consumers are inattentive to electricity prices, the costs of individual end-uses are not salient, and aggregate costs are only revealed intermittently. Consequently there is very little demand elasticity. Yet many utilities and policy makers would like to shift demand to periods of time when it is less privately or socially costly. Some utilities would like to meet energy efficiency and conservation goals by providing better information to their customers about their true costs. Furthermore economists argue that if greenhouse gases and other pollutants are to be efficiently controlled, buyers must know and pay the true marginal social cost of their energy uses. These hopes place a great onus on the ability of smart meter infrastructure to communicate prices in ways that will move demand.

The presence of habit persistence in consumption alters the amount and duration of demand response to price changes and new information. A large macroeconomic habit formation literature has shown that price changes that are perceived as permanent will have larger long-run impacts when consumption is governed by habits, but these effects will take time to accumulate as new steady state habit stocks are formed. On the other hand, households will be less sensitive to perceived transitory price changes than they otherwise would be. The impacts of new information, as opposed to price changes, on short and long run consumption with habits has not been thoroughly studied and we provide some preliminary empirical results. When designing new pricing and information programs enabled by the smart grid, utilities will need to consider how the effect on habits of their demand-shifting information and pricing programs.

5. Empirical Approach

Following Scott (2012), we estimate a habit persistence model using lags of consumption and leads of "prices". The first order conditions for an optimal consumption path with habit formation include the habit stock, which summarizes consumption lags, as well as the impact on all future consumption through future realizations of the habit stock. The current habit stock is captured using lags of consumption, and expected future consumption is captured using a set of price expectations at the current time period. We note that even in a durable habits model the importance of consumption and price lags and leads diminishes as the distance from the current period grows. We therefore truncate the number of lags and leads despite the possibility of habit persistence of long duration.

The approach of using price expectations is complicated by the fact that the rate structure (price) of electricity does not change frequently in California because of the regulatory process required to approve changes, and prices were fairly stable during the nine-month time frame we study[2]. However, the primary household service provided by electricity consumption in California is indoor temperature stabilization through heating and cooling. Therefore, the "effective price" of comfort achieved by maintaining a fixed indoor temperature varies closely with the outdoor temperature. We use hourly temperature observations to construct daily measures of "effective prices" of the heating and cooling services provided by electricity. A set of weather forecasts can therefore play the same role in the regression analysis as a set of price expectations.

We estimate models of the form

$$x_{it} = \beta_0 + \sum_{s=1}^{s'} \gamma_s x_{i,t-s} + \sum_{\tau=1}^{\tau'} \mathbf{T'}_{t+\tau}\beta_{T,\tau} + \sum_{\tau=1}^{\tau'} T_{h,t+\tau}\mathbf{W'}_i\beta_{h,\tau} + \sum_{\tau=1}^{\tau'} T_{c,t+\tau}\mathbf{W'}_i\beta_{c,\tau} + u_{iwm} + \epsilon_{it} \quad (20)$$

where x is the daily household-level consumption. $T_{h,t+\tau}$ and $T_{c,t+\tau}$ measure the number of hours on day $t+\tau$ that the outdoor temperature deviated from 65 degrees so that heating and cooling would be required; these are the "effective prices" described above and are referred to as "Heating and Cooling Degree Hours"[3]. $\mathbf{T}$ is a vector of quadratic polynomial terms on $T_{h,t+\tau}$ and $T_{c,t+\tau}$. $\mathbf{W}_i$ is a vector of household-specific covariates described in the next section, while the u's capture fixed effects at household-weekday-month level to capture unobserved household-specific effects that may vary by weekday and season. We estimate specifications with $s' = 7$ and $\tau' = 1$ to capture the potential for long memory, durable habits, as well as specifications with $s' = 1$ and $\tau' = 7$ to capture short memory habits.

We are aware of dynamic panel concerns with this estimation strategy - that the inclusion of lagged dependent variables with u_i introduces a bias in the coefficients by construction.

[2]The actual prices are listed in the appendix.

[3]Empirical electricity studies typically use monthly household data rather than daily, and transform raw temperature data into "Heating Degree Days" and "Cooling Degree Days". Our measure is exactly the same at a higher frequency.

Because of computational constraints, our dataset is too large to directly correct for this problem using system GMM. However, the dynamic panel bias is shrinking in the time dimension of the panels and our panels are fairly long - the average time dimension is 189 days. In other work (Gilbert and Graff Zivin (2013)) we estimate models that provide upper and lower bounds for the bias and show that the bias is small for this data, and coefficients are fairly stable whether we use one-way fixed effects at the household level or the household-weekday-month level, or two-way fixed effects at the household-weekday and household-month level.

6. Data

We use a sample of hourly electricity consumption observations and monthly billing information for approximately 10,000 households in Escondido, California, an inland suburb of San Diego. The data was provided by San Diego Gas & Electric (SDG&E) and was combined by the authors with hourly outdoor temperature data from weather stations in the area. The hourly observations span April, 2009 through January, 2010 while the monthly billing information covers the same time window and three additional years of historical bills. We use the billing data to discern the timing of the billing cycle and bill arrivals for each household. We included only households that had the same account-holder for the entire three-year billing history period in order to focus on households that are as close as possible to a steady state habit stock. We also included only households with at least seven months of daily or hourly data. The resulting data set consists of 10,943 households.

We use daily consumption as our dependent variable, and we use relationships between hourly consumption and hourly weather to construct household covariates for $\mathbf{W_i}$ that were unavailable from SDG&E. In order to do so we ran the following regression on each household:

$$\begin{aligned} x_{ih} &= \sum_{h=1}^{24} \phi_{ih}\Psi_h + \sum_{m=1}^{10} \beta_{im} Month + \sum_{d=1}^{7} \beta_{id} Weekday + \beta_{iC}\mathbb{1}\{Temp_h > 65\} \cdot (Temp_h - 65) \\ &\quad + \beta_{iH}\mathbb{1}\{Temp_h < 65\} \cdot (65 - Temp_h) + \nu_{ih} \quad \forall i \end{aligned} \tag{21}$$

with hourly dummies captured by Ψ_h. $\mathbb{1}$ takes a value of 1 if the temperature was above or below a 65-degree threshold. The household-specific coefficients and goodness of fit from these regressions capture household building and behavioral characteristics.

The variable "fit" is the R-squared from each regression; it summarizes the sensitivity of consumption to hourly and seasonal patterns, and weather variation. A tight relationship over hours within a given day implies larger air-conditioning and heating units. "Pattern" is the percentage difference between the maximum and minimum ϕ_{ih} for each household and captures daytime occupancy. The air-conditioning and heating capacity of the household are captured by β_{iC} and β_{iH}, which we rename "hotco" and "coldco", respectively, in the summary statistics table. Household-specific daily consumption volatility was also calculated in a variable called "dayonday". We also use the percentile rank in the sample of average monthly electricity consumption as a proxy for income. Summary statistics for variables used in the analysis are presented in Table 1. For more details, see Gilbert and Graff Zivin (2013).

Table 1: Summary Statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
daily kWh	20.724	18.354	0	628.810	2065072
monthly kWh	623.412	511.473	0	14310	2944589
bill ($)	114.897	127.405	5.07	2955.21	2944589
cdh65	106.822	112.644	0	503	2944589
hdh65	118.499	117.663	0	444	2944589
AC proxy	0.025	0.032	-0.103	0.276	2955529
heat proxy	-0.008	0.017	-0.173	0.191	2955529
fit	0.097	2.448	-27.625	30.705	2955529
pattern	-0.634	0.759	-16.871	39.947	2955529
dayonday	5.097	3.496	0.013	51.558	2955529

Unit roots in electricity usage have been found in many studies dating back to Engle, Granger and Hallman (1989). All studies that we know of that have found unit roots use regional aggregates over multiple years. Unit roots are to be expected over a longer time horizon because of permanent changes associated with population growth or migration, changes in energy-using durables and housing stock across the population, and policy changes that affect demand. Over a nine month period, it is less likely that we will see enough permanent innovations to the process to find unit roots in the daily series. We implement a Fisher-type panel augmented Dickey-Fuller test on the data and find that unit roots are strongly rejected in this data, as shown in table 2. The data was deseasonlized by household before implementing this test by regressing raw daily consumption separately for each household on monthly and weekday dummy variables and capturing the residuals from that regression. The panel of residual series was used in the ADF test.

This test conducts an ADF test on each panel separately and averages the p-values from each test using four methods: an inverse chi-squared, an inverse normal, an inverse logit, and

Table 2: Fisher-type ADF tests for unit roots in a panel

	Statistic	p-value
Inverse chi-squared(21886)	2.94e+05	0.0000
Inverse normal	-459.0191	0.0000
Inverse logit t(54719)	-774.2238	0.0000
Modified inv. chi-squared	1301.4227	0.0000
Panels	10943	
Avg. Periods	188.71	

Ho: All panels contain unit roots

Ha: At least one panel is stationary

a modified inverse chi-squared. A drawback to this test is that only one panel needs to be stationary in order to reject the null, at least for the inverse chi-squared version of the test. The other three methods of calculating the average p-value allow the number of stationary panels sufficient to reject the null to grow at the same rate that the number of panels tends toward infinity. The advantages of this test are that it allows unbalanced panels and missing data, which are two features that our data exhibit. We are unaware of other tests that allow both of these features but have the null hypothesis as all panels stationary which would be preferred to demonstrate that stationarity is the predominant feature of the data. As such, there may be some or even many individual households with unit roots, but we do not think this is the case.

7. Results

Results of the short-memory habit model are presented in table 3. Each column reports results from specifications including more covariates. Based on the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), the fourth column with all covariates is the preferred specification. All models show strong evidence of habit formation based on the statistically significant estimates of γ_1. A γ_1 of 0.42 (and far from 1) is consistent with panels being stationary on average but still exhibiting habit persistence in consumption.

Table 4 presents results that are similarly organized from the habits-as-durables model. Based on the AIC, BIC, and the long sequence of statistically signifcant and mostly positive coefficients on up to seven lags of consumption, we conclude that the long memory, durable habits model is more appropriate for the sample as a whole, and column 4 of table 4 is our preferred specification. This would suggest that household routines accumulate over time and elements of these routines strongly determine present consumption.

Table 3: Lag coefficients from short memory model

	(1)	(2)	(3)	(4)
γ_1	0.48***	0.48***	0.43***	0.42***
	(0.00318)	(0.00318)	(0.00343)	(0.00352)
Leads of T_h and T_c	YES	YES	YES	YES
Household covariates $\mathbf{W}_i$	NO	YES	YES	YES
Leads of T_h and T_c interacted with $\mathbf{W}_i$	NO	NO	YES	YES
Leads of T_h and T_c polynomial terms	NO	NO	NO	YES
N	1570888	1570888	1570888	1570888
R^2	0.295	0.295	0.336	0.344
aic	9833243.9	9833216.4	9739203.2	9720828.4
bic	9833440.2	9833523.1	9741055.5	9722938.4

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Coefficients on leads of temperature and household covariates not shown

Table 4: Lag coefficients from habits-as-durables model

	(1)	(2)	(3)	(4)
γ_1	0.46***	0.46***	0.44***	0.43***
	(0.00338)	(0.00338)	(0.00333)	(0.00335)
γ_2	0.089***	0.089***	0.086***	0.082***
	(0.00318)	(0.00318)	(0.00309)	(0.00308)
γ_3	0.016***	0.016***	0.014***	0.010***
	(0.00308)	(0.00308)	(0.00300)	(0.00299)
γ_4	-0.0039	-0.0038	-0.0097***	-0.011***
	(0.00301)	(0.00301)	(0.00294)	(0.00293)
γ_5	0.012***	0.012***	0.0044	0.0051*
	(0.00300)	(0.00299)	(0.00292)	(0.00291)
γ_6	0.026***	0.026***	0.024***	0.023***
	(0.00310)	(0.00310)	(0.00302)	(0.00301)
γ_7	-0.030***	-0.030***	-0.019***	-0.017***
	(0.00282)	(0.00282)	(0.00278)	(0.00278)
1 Lead of T_h and T_c	YES	YES	YES	YES
Household covariates $\mathbf{W}_i$	NO	YES	YES	YES
Lead of T_h and T_c interacted with $\mathbf{W}_i$	NO	NO	YES	YES
Lead of T_h and T_c polynomial terms	NO	NO	NO	YES
N	709532	709532	709532	709532
R^2	0.321	0.321	0.345	0.351
aic	4478831.7	4478644.2	4453140.3	4446241.0
bic	4478946.4	4478862.2	4453564.8	4446699.9

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Coefficients on lead of temperature and household covariates not shown

7.1. Habits and Salience

Having verified that habit persistence exists in electricity consumption and that the habit process likely has a long memory, we now explore the implications of this finding for the issue of intermittent price salience that our theoretical model addresses. We use the arrival of bills to investigate how households behave when they are reminded of an expenditure that is usually not salient. We argue that the response to bill arrivals will provide intuition about how households will respond to usage alerts and information campaigns promoted by utilities through the smart grid infrastucture.

In order to investigate this question we modify our estimating equations by adding interactions of lags of consumption with a dummy variable representing a window of time following the new billing cycle. We call this variable "post" in the following results tables. We do not observe exactly when households open the letter or email containing their bill, but we do know the dates that bills are sent. We define the window of time for the price reminder ("post") as the three-day period after the bill was sent, but we explore sensitivity to a seven-day window, a 10-day window, and a two-week window.

Table 5 gives the results of this exercise for the short-memory habits model. Although we believe that the long-memory process is better supported by the data, the results of the short memory model help provide intuition for interpreting the results. Recall from the model that ρ is the parameter governing how quickly the habit stock accumulates from more recent consumption. A larger ρ means that an agent forms new habits more quickly. A positive coefficient on the interaction term in the regression implies an increase in ρ during the period following the price reminder. In other words, table 5 indicates that receiving a price reminder causes households to show more urgency in trying to form new habits. The significance of this effect is eliminated when the price reminder window is defined over a broader time span, suggesting that ρ is returning to its typical value and households eventually put less effort into forming new habits. The negative coefficients on the "post" variable also indicate that electricity use was reduced during the reminder window, providing evidence of intermittent price salience that is consistent with our model and the results in (Gilbert and Graff Zivin, 2013).

Extending this analysis to results from the habits-as-durables model in table 6, we see that the largest impacts of the price reminder on habits occur when the reminder period is defined as three days, and operate on the most recent lags of consumption. However, as the window of time is broadened, significant positive coefficients are found on the interactions with further lags, suggesting that the weight placed on past consumption is being shifted further away from new habits as more time passes after the price reminder arrives.

Table 5: The effect of a price update on short memory habits

	3 days	7 days	10 days	14 days
post	-0.49***	-0.37***	-0.40***	-0.14***
	(0.0157)	(0.0120)	(0.0112)	(0.00937)
γ_1	0.42***	0.42***	0.41***	0.41***
	(0.00373)	(0.00367)	(0.00275)	(0.00309)
post*γ_1	0.028***	-0.0080	0.0087	0.0084
	(0.00963)	(0.00924)	(0.00972)	(0.00703)
N	1570888	1570888	1570888	1570888
R^2	0.344	0.344	0.345	0.344
aic	9719634.8	9719709.6	9719052.8	9720555.3
bic	9721769.3	9721844.0	9721187.3	9722689.8

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Leads of temperature, polynomial terms, and interactions included in all columns

Table 6: The effect of a price update on durable habits

	3 days	7 days	10 days	14 days
post	-0.40***	-0.28***	-0.41***	-0.21***
	(0.0263)	(0.0190)	(0.0169)	(0.0151)
γ_1	0.43***	0.43***	0.43***	0.42***
	(0.00347)	(0.00372)	(0.00392)	(0.00412)
γ_2	0.076***	0.074***	0.075***	0.071***
	(0.00320)	(0.00345)	(0.00363)	(0.00399)
γ_3	0.0088***	0.0068**	0.0061*	0.0049
	(0.00309)	(0.00330)	(0.00346)	(0.00377)
γ_4	-0.011***	-0.015***	-0.016***	-0.010***
	(0.00306)	(0.00326)	(0.00336)	(0.00366)
γ_5	0.0050*	0.0032	-0.00057	0.00014
	(0.00303)	(0.00324)	(0.00344)	(0.00364)
γ_6	0.024***	0.020***	0.013***	0.0073*
	(0.00309)	(0.00331)	(0.00353)	(0.00384)
γ_7	-0.020***	-0.021***	-0.023***	-0.020***
	(0.00285)	(0.00306)	(0.00327)	(0.00368)
post*γ_1	0.047***	0.014*	0.010	0.032***
	(0.0128)	(0.00820)	(0.00721)	(0.00678)
post*γ_2	0.071***	0.040***	0.023***	0.026***
	(0.0118)	(0.00758)	(0.00680)	(0.00626)
post*γ_3	0.013	0.016**	0.011	0.013**
	(0.0120)	(0.00782)	(0.00695)	(0.00619)
post*γ_4	-0.012	0.011	0.0096	-0.0028
	(0.0101)	(0.00732)	(0.00681)	(0.00607)
post*γ_5	-0.0055	-0.0011	0.011*	0.011*
	(0.0112)	(0.00750)	(0.00652)	(0.00604)
post*γ_6	-0.016	0.011	0.028***	0.034***
	(0.0124)	(0.00809)	(0.00682)	(0.00624)
post*γ_7	0.026**	0.017**	0.012**	-0.0015
	(0.0119)	(0.00747)	(0.00634)	(0.00572)
N	709532	709532	709532	709532
R^2	0.353	0.352	0.353	0.353
aic	4444699.0	4445006.4	4444442.0	4444718.5
bic	4445249.6	4445557.1	4444992.6	4445269.2

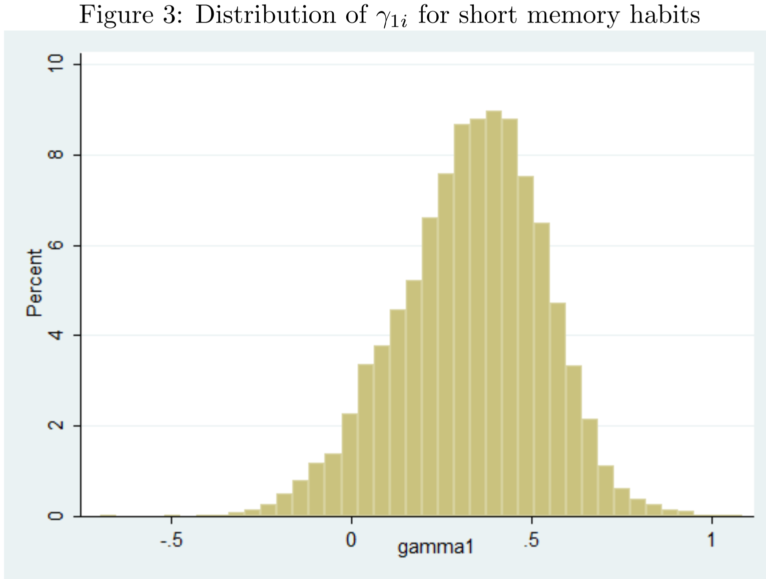
Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

One lead of temperature, polynomial terms, and interactions included in all columns

Lastly, we investigate heterogeneity in habit persistence between households by estimating the short- and long-memory models on each household separately. For the short memory model, we display the distribution of estimated γ_1 across households in the figure below. Seventy-four percent of these are positive and statistically significant, suggesting that some degree of habit formation is a dominant feature of this consumption good, although there is a significant minority of households whose consumption series is not well-described by this model. For the long-memory model, the distribution of γ_1 is very similar to the distribution from the short memory model. When we conduct an F-test of the second through seventh lags, however, in order to find out how many households have a long memory habit process, we find that only 18% of households have a p-value less than 5%. The long memory process may therefore be a strong feature of only a subset of households. Clearly the correct habit duration varies by household and may lie between one lag and seven lags, so 18% is a lower bound on the share of households with long memory habits. Further investigation of this heterogeneity is a topic of ongoing research. Identifying households with different types of habit processes suggests that alternative programs may need to be designed for different household types.

Figure 3: Distribution of γ_{1i} for short memory habits



8. Conclusion

This is the first study that we know of to demonstrate habit persistence in daily consumption. We find support for a habits-as-durables model suggesting that even distant past consumption influences present consumption. The context - residential electricity use - gives rise to important behavioral and policy implications. Insalient prices or inattention to prices can lead to inefficiently large habit stock accumulation. We find some evidence that information or cognitive constraints are leading to excess consumption because households engage in temporary conservation in the days after receiving a bill even though true marginal prices have not changed. The effect of information on habit formation is an important topic for future research.

These stocks make it more difficult to achieve short-term demand responses on days when electricity grids are constrained and peak power is expensive to deliver. This finding does provide some optimism for the long-term effectiveness of more permanent pricing policies like pollution taxes. However, habits are heterogeneous across households and it may be the case that a minority of households account for the predominance of long memory habits. Exploring the patterns of this heterogeneity and defining the implications for Pigouvian taxes and peak pricing demand response programs is an important topic for future research.

Disclosure Statement

Ben Gilbert was employed by Sempra Energy, the parent company of San Diego Gas & Electric, within three years prior to the beginning of this study. This employment relationship ended before data was obtained and analyzed for this project, and the duties of employment were unrelated to the study presented here. Ben Gilbert has maintained no financial relationship with Sempra or its subsidiaries over the duration of this study or subsequently.

References

Allcott, H. (2011). Social norms and energy conservation. Journal of Public Economics. 95 (9): 1082-1095.

Allcott, H. and S. Mullainathan (2010). Behavior and energy policy. Science, 327 (5970): 1204-1205.

Allcott, H., Mullainathan S., & Taubinsky, D. (2014). Energy Policy with Externalities and Internalities. Journal of Public Economics, Forthcoming.

Allcott, H. and Rogers, T. (2012). The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. (No. w18492). National Bureau of Economic Research.

Becker, G. S., Grossman, M., & Murphy, K. M. (1994). An Empirical Analysis of Cigarette Addiction. The American Economic Review, 84(3), 396-418.

Browning, M. (1991). A simple nonadditive preference structure for models of household behavior over time. Journal of Political Economy, 607-637.

Chetty, R., Looney, A., & Kroft, K. (2007). Salience and taxation: Theory and evidence (No. w13330). National Bureau of Economic Research.

Darby, S. (2006). The effectiveness of feedback on energy consumption: a review for DEFRA of the literature on metering, billing and direct displays," Technical Report, Environmental Change Institute.

DellaVigna, S. (2009). Psychology and economics: Evidence from the field. Journal of Economic Literature, 315-372.

Engle, R. F., Granger, C. W., & Hallman, J. J. (1989). Merging short-and long-run forecasts: An application of seasonal cointegration to monthly electricity sales forecasting. Journal of Econometrics, 40(1), 45-62.

Finkelstein, A. (2009). E-ztax: Tax salience and tax rates. The Quarterly Journal of Economics, 124(3), 969-1010.

Gilbert, B. & Graff Zivin, J. (2013). Dynamic Salience with Intermittent Billing: Evidence from Smart Electricity Meters. (No. w19510). National Bureau of Economic Research.

Graff Zivin, J., Kotchen, M., & Mansur, E. (2012). Spatial and Temporal Heterogeneity of Marginal Emissions: Implications for Electric Cars and Other Electricity-Shifting Policies. (No. w18462). National Bureau of Economic Research.

Grubb, M., & Osborne, M. (2012). Cellular service demand: Biased beliefs, learning, and bill shock.

Hall, R. E. (1978). Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence. The Journal of Political Economy, 86(6), 971-987.

Heien, D., & Durham, C. (1991). A test of the habit formation hypothesis using household data. The Review of Economics and Statistics, 189-199.

Hsiaw, A., & Harding, M. (2013). Goal setting and energy efficiency. Working Paper.

Jacobsen, G. D., Kotchen, M. J., & Vandenbergh, M. P. (2012). The behavioral response to voluntary provision of an environmental public good: Evidence from residential electricity demand. European Economic Review, 56(5), 946-960.

Jaffe, Adam B. and Robert N. Stavins (1994). The energy paradox and the diffusion of conservation technology. Resource and Energy Economics, 5, 16 (2), 91-122.

Jessoe, Katrina and David Rapson (2014). Knowledge is (Less) Power: Experimental

Evidence from Residential Energy Use. Forthcoming, American Economic Review.

Karlan, D., McConnell, M., Mullainathan, S., & Zinman, J. (2010). Getting to the top of mind: How reminders increase saving (No. w16205). National Bureau of Economic Research.

Leth-Petersen, S. (2007). Habit formation and consumption of energy for heating: evidence from a panel of Danish households. The Energy Journal, 35-54.

McKinsey Global Institute. "Wasted Energy: How the U.S. Can Reach its Energy Productivity Potential,", Technical Report, June 2007.

Pollak, R. A. (1970). Habit formation and dynamic demand functions. The Journal of Political Economy, 1970, 78 (4), 745-763.

Scott, K. R. (2012). Rational habits in gasoline demand. Energy Economics.

Sexauer, B. (1977). The role of habits and stocks in consumer expenditure. The Quarterly Journal of Economics, 127-142.

Sexton, S. E. (2011). Automatic Bill Payment, Price Salience, and Consumption: Evidence from Residential Electricity Consumption. Working Paper.

Shah, A. K., Mullainathan, S., & Shafir, E. (2012). Some consequences of having too little. Science, 338(6107), 682-685.

Spinnewyn, F. (1981). Rational habit formation. European Economic Review, 15(1), 91-109.

Stango, V., & Zinman, J. (2011). Limited and varying consumer attention: Evidence from shocks to the salience of bank overdraft fees (No. w17028). National Bureau of Economic Research.

Taubinsky, D. (2013). From Intentions to Actions: A Model and Experimental Evidence of Inattentive Choice. Working Paper.

Appendix A. Tables

Table A.7: SDG&E Rate Structures 4/15/2009 to 1/31/2009

Dates	Rates ($/kWh)	DR	DRLI	A (flat rate)
4/15/2009	Baseline (11.5 kWh/day)	0.04363	0.04854	0.06544
	101% to 130% of Baseline	0.0638	0.06871	
	131% to 200% of Baseline	0.21818	0.16067	
	Over 200% of Baseline	0.23818	0.16067	
5/1/2009	Baseline (11.8 kWh/day)	0.01692	0.02183	0.07508
	101% to 130% of Baseline	0.03709	0.042	
	131% to 200% of Baseline	0.20379	0.14472	
	Over 200% of Baseline	0.22379	0.14472	
9/1/2009	Baseline (11.8 kWh/day)	0.01692	0.02183	0.07239
	101% to 130% of Baseline	0.03709	0.042	
	131% to 200% of Baseline	0.1988	0.13973	
	Over 200% of Baseline	0.2188	0.13973	
11/1/2009	Baseline (11.5 kWh/day)	0.04455	0.04946	0.06253
	101% to 130% of Baseline	0.06472	0.06963	
	131% to 200% of Baseline	0.21019	0.15268	
	Over 200% of Baseline	0.23019	0.15268	
1/1/2010	Baseline (10.8 kWh/day)	0.06026	0.05867	0.07222
	101% to 130% of Baseline	0.08103	0.07944	
	131% to 200% of Baseline	0.18952	0.1394	
	Over 200% of Baseline	0.20952	0.1394	

Notes: During the time span our data covers (April 15, 2009 to January 31, 2010), residential electricity rates changed several times by small amounts. This table describes in detail the rates faced by our sample of households.