Efficient Retail Pricing in Electricity and Natural Gas Markets:

A Familiar Problem with New Challenges

Steven L. Puller[1] and Jeremy West[2]

January 2013

Abstract

A long line of research investigates whether the retail prices of electricity and natural gas send proper signals about scarcity in order to induce efficient consumption. Historically, regulated utilities have not designed tariffs that set marginal prices equal to marginal costs. Currently, some jurisdictions are opening the retail sectors of the gas and electricity industry to competition via "retail choice". These new regimes replace imperfect regulation with imperfect competition as the process by which retail tariffs are formed. We discuss the challenges in evaluating the efficiency of these new pricing regimes and present descriptive evidence of how pricing has changed in markets with retail choice.

---

[1] Texas A&M University and NBER. Corresponding author. Email: puller@econmail.tamu.edu
[2] Texas A&M University. Email: jwest@econ.tamu.edu.
We thank Koichiro Ito for useful discussions and Kim Castanon and Robert Manning for research assistance.

1. The Pricing Problem in Retail Electricity and Natural Gas

Researchers have investigated the efficiency of retail pricing in the electricity and natural gas industries for decades. Historically, the challenge has been pricing in a manner that ensures that a regulated utility covers its investment and production costs while simultaneously providing consumers with optimal incentives to consume. Although some jurisdictions continue to follow this regulated utility model, others -- most notably in the U.S. and UK -- are increasingly opening retail electricity and natural gas sectors for "competition." Obviously, the retail tariffs that arise from a market can differ from those that are the outcome of a political and regulatory process.

Despite regulatory restructuring, these industries still face some of the same pricing problems as in the past. In the electricity industry, the transmission and distribution (T&D) sectors have natural monopoly characteristics – fixed costs make it very inefficient to have multiple sets of long-distance transmission lines or local distribution infrastructure (e.g. substations, transformers, and power lines). Local distribution in the natural gas industry also has natural monopoly characteristics of high fixed costs and low variable costs. To recover these large transmission and distribution costs and ensure adequate investment, the T&D components must be "priced" into retail electricity and natural gas.

The electricity industry faces another dimension along which costs need to be incorporated into pricing – the timing of consumption. Although natural gas can be stored (to an extent) in the distribution system, electricity cannot be efficiently stored on a large scale, which requires balancing production and consumption at every moment in time. Because demand follows strong daily patterns, the variable cost of production is often substantially higher during peak as compared to off-peak periods. Retail prices could vary as frequently as every ten minutes to reflect the marginal cost of production at the customer's location in the current state of the electric distribution grid. Joskow and Wolfram (2012) discuss some of the current challenges in dynamic pricing. For residential customers, there is currently little take-up of time-varying prices, but Borenstein (2012) discusses how such pricing schemes can be designed to facilitate take-up and minimize adverse distributional consequences. Larger commercial and industry customers exhibit more widespread use of dynamic pricing, which we show evidence of below.

Throughout most of the U.S. experience, the distribution sector of the natural gas industry and all sectors of the electricity industry have been regulated utilities. In such environments, the standard challenge from an efficiency perspective is to set prices in a manner that ensures adequate recovery of investment costs while at the same time providing customers with optimal consumption incentives on the margin. A rich theoretical literature provides solutions to this pricing problem. In the natural gas industry, one normative prescription arising from this literature is that first-best can be achieved by implementing a uniform two-part tariff – that is, a fee structure that requires all customers to pay a fixed "connection" charge that covers each customer's share of the fixed local distribution costs along with a marginal price that is set equal to marginal cost.

Because power is not storable in the electricity sector, the efficiency properties of different pricing schemes are more nuanced. Joskow and Tirole (2006) provide a detailed analysis of the efficiency of pricing under various scenarios for retail competition in electricity. First-best solutions are generally achieved only if retail customers have their consumption metered on a real-time basis and face retail prices reflecting real-time wholesale scarcity. In the absence of real-time meters, a second best outcome is achieved if a monopoly retailer uses two-part tariffs for its customers, who have their expected real-time consumption estimated with load profiles. Because the variable usage price does not adjust to match system grid conditions, customers over-consume during peak periods and under-consume off-peak. However, retail competition adds another source of inefficiency – a retailer is charged for wholesale power based upon the load profiled estimated demand of its customers rather than for actual demand, which is a function of the actual retail price. Finally, another source of inefficiency is imperfect competition. Joskow and Tirole (2006) note that their theoretical analysis assumes a competitive homogeneous good market, but Hortacsu, Mandanizadeh, and Puller (2012) show that at least the Texas retail market exhibits considerable consumer inertia due to search costs and perceived product differentiation by the incumbent retailer.

2. The Challenge of Practically Defining a Benchmark for Efficient Pricing

Despite fairly clear normative theoretical prescriptions, constructing a benchmark for first-best pricing is not straightforward in practice. We discuss several complicating factors below and show how difficult it is to define a benchmark or even to establish if marginal prices are "too high" or "too low".

One well-known complication of two-part tariffs is that charging a heterogeneous set of customers a uniform fixed charge may discourage consumption on the extensive margin – some "low value" consumers may choose not to consume the good at all if the fixed charge is larger than gross consumer surplus (Auerbach and Pellachio, 1979). In energy settings, this situation is unlikely for electricity consumption, but natural gas customers may opt to use other energy commodities rather than pay the "connection" charge for natural gas. For example, households may choose to heat their homes with less energy efficient fuels. This possibility of "disconnection" suggests that consumers be charged different fixed charges, or if that is not possible, that other second-best pricing strategies be employed.[3]

A second complication is that even if marginal prices are set equal to the marginal private costs of the electric or gas utility, there may be external costs that are not incorporated into the price signals. Production and consumption in the natural gas and electricity industries involve emission costs – both "local" (e.g. SO_2 and NO_x) and "global" ($C0_2$ and other greenhouse gases) – that are not internalized. In a first-best world in which all other goods are priced at marginal social costs, the socially optimal price of retail gas or electricity is above the marginal private costs that would arise from a utility employing the optimal two-part tariff. In a second best setting in which multiple substitute goods are "mispriced", it is generally difficult to assess whether the retail price of any single energy commodity should rise to reflect externalities, as "correcting" one distortion could enhance other distortions. Nevertheless, policy recommendations targeting a specific industry sometimes make the implicit assumption that increasing the price of an individual energy source to reflect the marginal social cost improves social welfare.

The nature of consumer price responsiveness serves as a third factor hindering policy recommendations from even suggesting the direction that prices should move to increase efficiency. Recent research investigates the possibility that utility consumers may not actually

[3] When "disconnection" is possible, Ng and Weisser (1974) show an application of the theory of second best that pricing below marginal cost can be optimal.

respond to the *marginal* price of gas or electricity in the manner implied by standard theory. Electric and gas utilities typically charge non-linear tariffs including a fixed "customer charge" along with variable "volumetric" charges, which vary in usage. These volumetric charges often take the form of increasing block tariffs.

Suppose that customers respond to larger utility bills by consuming less energy, ceteris paribus. This observed behavior is consistent with consumers responding to either the *marginal* price (i.e. the slope of the non-linear tariff) or to the *average* price (e.g. the average ¢/kWh that is reported on some bills).[4] Either behavioral response is possible depending on consumer knowledge of the tariff function and the type of information that is saliently reported on customer bills. Utility bills often do a poor job of clearly displaying the marginal price of an additional kWh of power or Ccf of natural gas. Casual empiricism suggests that utility customers are better informed about the *total* monthly expenditures on gas/electricity rather than the marginal price; contrast this with retail gasoline customers who are likely better informed about the (marginal) price per gallon than about their total monthly expenditures on gasoline.

In general, average prices might be above or below marginal prices depending on the size of the fixed charge and the extent to which increasing block tariffs are utilized. If consumers respond to the average price rather than the marginal price, then they may either over or under-consume relative to the optimal level of consumption as implied by standard theory. Ito (2012) tests how consumers respond to nonlinear tariffs by using a spatial discontinuity research design, which exploits variation across two contiguous electric utilities employing different tariff

[4] If the level of consumption is uncertain due to monthly shocks in the demand for energy services, consumer may respond to the *expected* marginal price (Borenstein, 2009).

functions. His results suggest that consumers respond to changes in average prices rather than to marginal or expected marginal prices.

This line of research suggests that if normative prescriptions from academic research focus on "getting the marginal price right," then they should also advocate for bill design that saliently displays this price signal. There is no point in focusing attention on one price signal if consumers respond to another signal.

3. Retail Competition and the Efficiency of Pricing in the U.S.

During the last decade, the number of retail customers purchasing energy services from a firm other than a regulated utility increased substantially. Several states have opened retail markets in which non-incumbent firms procure wholesale power and gas and market it to commercial, industrial, and residential customers. In the U.S. this phenomenon is more prevalent in electricity than natural gas, so we will focus our discussion of retail competition on electricity.

States continue to move from regulated regimes to competitive retail electricity markets, which generates the important research question of whether retail utility competition increases pricing efficiency.   Two dimensions of pricing have particularly high potential to improve welfare. Probably the most fruitful adjustment would be to increase the implementation of dynamic pricing across all classes of customers. The take-up among residential customers is currently very low, although we show some evidence below suggesting greater promise among large commercial and industrial customers.

A second feature of retail competition is that the tariffs arising from competitive firms could be more efficient than those determined by a regulatory process. Although time-varying prices are necessary to achieve the first-best solution, it remains unlikely in the near-term that many residential customers will face time-varying tariffs instead of charges based solely on total monthly consumption. Therefore, it is valuable to investigate how well monthly tariffs reflect the marginal cost of power (averaged over the hours of the month) and generate proper (second-best) consumption signals. We discuss initial evidence on both of these issues using data from Texas, the state which ranks highest in electricity consumption and has a high rate of participation in retail electricity choice programs.

A. Prevalence of Time-varying Prices for Commercial and Industrial Customers

Retail competition changes the nature by which tariffs are formed for commercial and industrial (C&I) customers. In some jurisdictions with retail choice, C&I customers bilaterally negotiate contracts for power from retail providers. If these large customers have meters recording consumption at fine intervals, it is possible for contracts to specify prices that vary with the consumption during hours when wholesale prices vary. Because the tariffs in these jurisdictions arise from individual contracting rather than a political rate-making process, the resulting tariffs could differ. For example, the tariffs in markets with retail choice could involve more time-varying price contracts either because the bilaterally negotiated tariffs better share risk or because they are less likely to involve cross-subsidies between customers with different daily load profiles.[5]

[5] For a discussion of cross-subsidies and the design of real-time pricing programs, see Borenstein (2007).

We present descriptive evidence that notable numbers of large C&I customers in Texas are signing bilateral contracts that subject the end-user to time-varying prices. In April 2009, the Texas grid operator, ERCOT, requested that all retailers identify whether each customer's contract provided incentives to reduce consumption when spot prices were high. ERCOT provided us with data for the 8,537 C&I customers that are metered with interval data recorders. These customers represent approximately 20% of all consumption in ERCOT. The data include one year of interval-level consumption data and an indicator variable for whether the retail contract provided "a financial incentive or requirement to reduce consumption in response to high wholesale spot prices." The survey specifies that this definition is intended to include real-time pricing and critical-peak pricing but to exclude time-of-use pricing. See Cancho and Puller (2012) for further details and an analysis of the price elasticity of demand. These data indicate that 15% of these large C&I customers have contracts with incentives to respond to hourly spot prices. The customers on time-varying prices tend to be larger consumers – the consumption-weighted fraction of customers subject to time-varying prices is 30%. This does not necessarily imply that all these customers' load is subject to time-varying prices, but it does mean that on the margin these customers face price signals of real-time scarcity. This is suggestive evidence that jurisdictions with retail choice for large customers may develop meaningful demand response programs. Of course, the observed take-up of time-varying prices could result from the nature of the retail market or other factors affecting the industry as a whole.

B. Efficiency of Monthly Tariffs to Residential Customers

As we note above, time-varying prices are the most efficient means to send price signals, but dynamic pricing schemes may not get traction among residential customers, especially in the near term. Therefore, we explore whether (second-best) monthly tariffs send accurate signals about the (monthly average) marginal cost of power. It is a priori ambiguous whether transitioning to competition would drive marginal retail price to marginal cost – tariffs with the inefficiencies generated by the regulatory process are replaced by those emerging from imperfectly competitive retail markets.

The outcome of the regulatory process is known to generate multi-part tariffs in which the marginal price diverges from marginal cost. Fixed distribution costs are priced into the variable/usage component of tariffs rather than solely into the fixed customer charge. For the case of natural gas, Davis and Muehlegger (2010) find that residential and commercial customers face retail prices that are more than 40% above marginal cost. Also, state public utility commissions often establish increasing block rates so that higher usage consumers pay a higher price on the margin. These increasing block tariffs are poorly conceived from an efficiency perspective. One motivation is that higher marginal prices encourage conservation; however, the marginal social cost of one more unit of consumption is not any higher for an individual consumer's 499th kilowatt-hour than the consumer's 500th kWh. Equity concerns often serve as another motivation, but Borenstein and Davis (2010) find only a weak correlation of natural gas consumption with income. Finally, under current net energy metering policy, transmission and distribution costs will continue to inflate usage charges as customers with rooftop solar and other distributed generation effectively avoid some transmission and distribution costs.

Under a competitive retail regime, other market forces influence tariff design. Joskow and Tirole (2006) show that a perfectly competitive retail market theoretically yields linear pricing of energy sales (grid access charges are assumed to be zero). In an imperfectly competitive retail market with consumer inertia, the tariff set by the incumbent firm, perhaps via regulatory mandate, is likely to impact the tariff functions chosen by the competitive retailers.

To provide suggestive evidence of the shift in tariffs associated with the introduction of retail competition, we show how the retail price-cost margins for markets opened to competition change as compared to those that remain regulated. In Texas, areas of the state served by investor-owned utilities were required to allow entry of competitive retailers beginning in January of 2002. The incumbent retailer charged a regulated tariff called the price-to-beat, and new entrant retailers were allowed to choose any tariff functions.[6] Areas served by municipal utilities were allowed to opt-out of retail choice, and all Texas munis did so. Our analysis follows the spirit of a difference-in-differences approach; however, we do not have sufficient data to perform formal inference. Thus, we view our evidence as merely descriptive analysis of price changes following a particular introduction of retail electricity competition. We measure the extent to which marginal prices differed from marginal cost before and after retail competition began in 2002. Our approach is similar in spirit to Davis and Muehlegger (2010); the primary differences are that we focus only on one state and use actual tariff functions whereas Davis and Muehlegger study nationwide pricing but use estimates of average tariff functions.

[6] For more details of retail choice in Texas, see Hortacsu, Madanizadeh, and Puller (2012).

We collected data on the monthly residential retail tariff functions pre- and post-2002—specifically, from January 1997 through July 2006--for one large territory that deregulated its retail market (Reliant Energy serving the greater Houston area) and two large territories that did not deregulate (greater San Antonio and Austin). In addition, we calculated the volume-weighted-average monthly wholesale price of power in ERCOT, matching this to the retail territories by region. This allows us to measure the monthly difference between the actual *marginal* retail price (i.e. the slope of the tariff function) and realized marginal energy costs for any usage level in each territory. In presentation of our results below, we consider this difference for the usage level corresponding to the median household consumption for each month of the year.[7]

Figure 1 displays the differences between marginal price and marginal cost before and after retail "deregulation" in 2002. The "Control" series is the average retail markup across the two areas that never allowed retail choice; the average margin was 5.02¢/kWh prior to 2002 and 3.92¢/kWh after 2002. There are three "Treatment" series that are informative about the changes under retail choice. The "Incumbent" was required to charge the regulated price-to-beat, so the incumbent's pricing does not represent the "outcome" of retail choice but rather is a driver of market forces under retail choice. For the incumbent, the average margin was 6.77¢/kWh prior to 2002 and 7.49¢/kWh after 2002. The series "Treated Large Entrant" shows that a major new entrant priced only slightly lower than the price-to-beat (6.38¢/kWh on average). The series "Treated Lowest Cost Entrant" plots the retail price margin (3.49¢/kWh on average) for the

[7] The usage distribution is calculated using a large sample of residential consumers spread geographically across Texas for the period 2002-2006. We view the median consumption level in each calendar month as a reasonable proxy for the consumption pattern of a "representative consumer."

entrant firm that each month yielded the lowest total bill for the median consumption level. We do not have information on market shares for individual entrants but the incumbent continued to serve over half the market throughout the sample. This series shows margins that are substantially lower than the large entrant's margin and are very close to margins for the municipal "control" areas. The upshot of this analysis is that the marginal retail price diverged from marginal cost both before and after retail deregulation.

We also explore the extent to which increasing block tariffs continue to be used under a retail choice regime. As we discuss above, tiered pricing schemes are inconsistent with marginal cost pricing (either social or private), and often are motivated by distributional or conservation goals (Borenstein and Davis, 2010). It is possible that tariffs formed instead via market forces could be linear tariffs for all usage charges and show no differential marginal prices. However, evidence from the first years of the Texas retail market suggests that increasing block tariffs continue under retail choice. For each month and provider, we compute the marginal usage price across a distribution of usage levels and then calculate the standard deviation of these marginal prices. If a provider levies only a single block tariff for a given month, then all customers pay the same marginal price and the standard deviation is zero. However, if the provider uses an increasing (or decreasing, as is occasionally the case) block tariff, customers pay different marginal prices and the standard deviation of these prices will depend upon the size of price difference between the tiers and the frequency with which customers consume on the different tiers.

Figure 2 plots the monthly standard deviation for several providers before and after retail choice was introduced in 2002. The "Control" areas that never had retail choice display no systematic difference in standard deviation pre- and post-2002. The series for the "Treated Incumbent" indicates that the regulated price-to-beat utilized increasing block pricing that yielded similar levels of variation in marginal prices as prior to retail choice. The "Treated Large Entrant" series indicates that this retailer entered the market with single block tariffs for the first year but then switched to an increasing block tariff yielding similar levels of variation in marginal prices as the price-to-beat. (Interestingly, this entrant is an incumbent from another part of the state, and this firm used increasing block tariffs in its home territory in 2002 while introducing a single block tariff in the Houston territory). The final series "Treated Lowest Cost Entrant" displays the variation in marginal prices for the entrant retailer that delivered the observed consumption at lowest total expenditure. The identity of this lowest cost firm changed from month to month and across consumption levels within a month; however, the market share of these firms was likely very low. Therefore, this series is more informative about the tariffs offered to attract new customers than the tariffs that customers were actually paying. Nevertheless, even these tariffs included substantial variation in effective marginal prices.

Finally, let us comment on external validity. The changes in tariffs after "retail deregulation" can vary substantially across jurisdictions depending on how the transition to competition is regulated. States have taken different approaches on the rate charged by the incumbent firm, which in a market with consumer inertia can have a large impact on pricing by new entrant retailers. In Texas, the public utility commission regulated the incumbent's rate at a level that was thought to be above competitive levels with the goal of encouraging entry by other retailers. States that follow different transition strategies could experience different changes in the shape and level of tariff functions.

4. Concluding Remarks

The introduction of retail competition in electricity and natural gas markets generates a new dimension of the familiar question on the efficiency of pricing. Imperfect competition is replacing imperfect regulation, so an important new area of research is to study the efficiency of these new markets as they are created and mature.

Economists have focused attention on the marginal price as the relevant signal of scarcity. The extent to which consumers respond to this signal depends upon the saliency of information conveyed to them. In new retail markets, bills will be designed by firms who compete for customers rather than designed by legal counsel from regulatory commissions. Therefore, it will be interesting to observe how bill design changes under retail choice and which types of information on price and usage are saliently displayed on those bills. The type of information conveyed to consumers also will depend upon how many consumers utilize online or automatic bill pay with their retailers. Finally, research should continue to explore the competitiveness of these new retail markets and the inertia surrounding incumbent firms. Ultimately, the welfare implications for retail choice will depend upon the competitiveness of the retail markets, the nature of the tariff functions, the information about those tariff functions that is saliently conveyed to customers, and how the behavioral response of these factors interacts with other distortions in energy markets.

References

Auerbach, A. J. & Pellechio, A. J. (1978), 'The Two-Part Tariff and Voluntary Market Participation', *The Quarterly Journal of Economics* **92**(4), 571-587.

Borenstein, S. (2007), ‘Wealth Transfers among Large Customers from Implementing Real-Time Retail Electricity Pricing’, *Energy Journal* 28(2): 131-49.

Borenstein, S. (2009), ‘To What Electricity Price Do Consumers Respond? Residential Demand Elasticity Under Increasing-Block Pricing’, Working Paper.

Borenstein, S. (2012), 'Effective and Equitable Adoption of Opt-in Residential Dynamic Electricity Pricing', National Bureau of Economic Research Working Paper 18037.

Borenstein, S. & Davis, L. W. (2010), 'The Equity and Efficiency of Two-Part Tariffs in U.S. Natural Gas Markets', National Bureau of Economic Research Working Paper 16653.

Cancho, C. & Puller, S. L. (2012), 'Who responds to electricity prices? New evidence on demand response by large customers', Working Paper.

Davis, L. W. & Muehlegger, E. (2010), 'Do Americans consume too little natural gas? An empirical test of marginal cost pricing', *The RAND Journal of Economics* **41**(4), 791-810.

Hortaçsu, A.; Madanizadeh, S. A. & Puller, S. L. (2012), 'Power to Choose? An Analysis of Choice Frictions in the Residential Electricity Market', Working Paper.

Ito, K. (2012), 'Do Consumers Respond to Marginal or Average Price? Evidence from Nonlinear Electricity Pricing', Working Paper.

Joskow, P. & Tirole, J. (2006), 'Retail Electricity Competition', *The RAND Journal of Economics* **37**(4), 799-815.

Ng, Y.-K. & Weisser, M. (1974), 'Optimal Pricing with a Budget Constraint: The Case of the Two-part Tariff', *The Review of Economic Studies* **41**(3), 337-345.

Figure 1

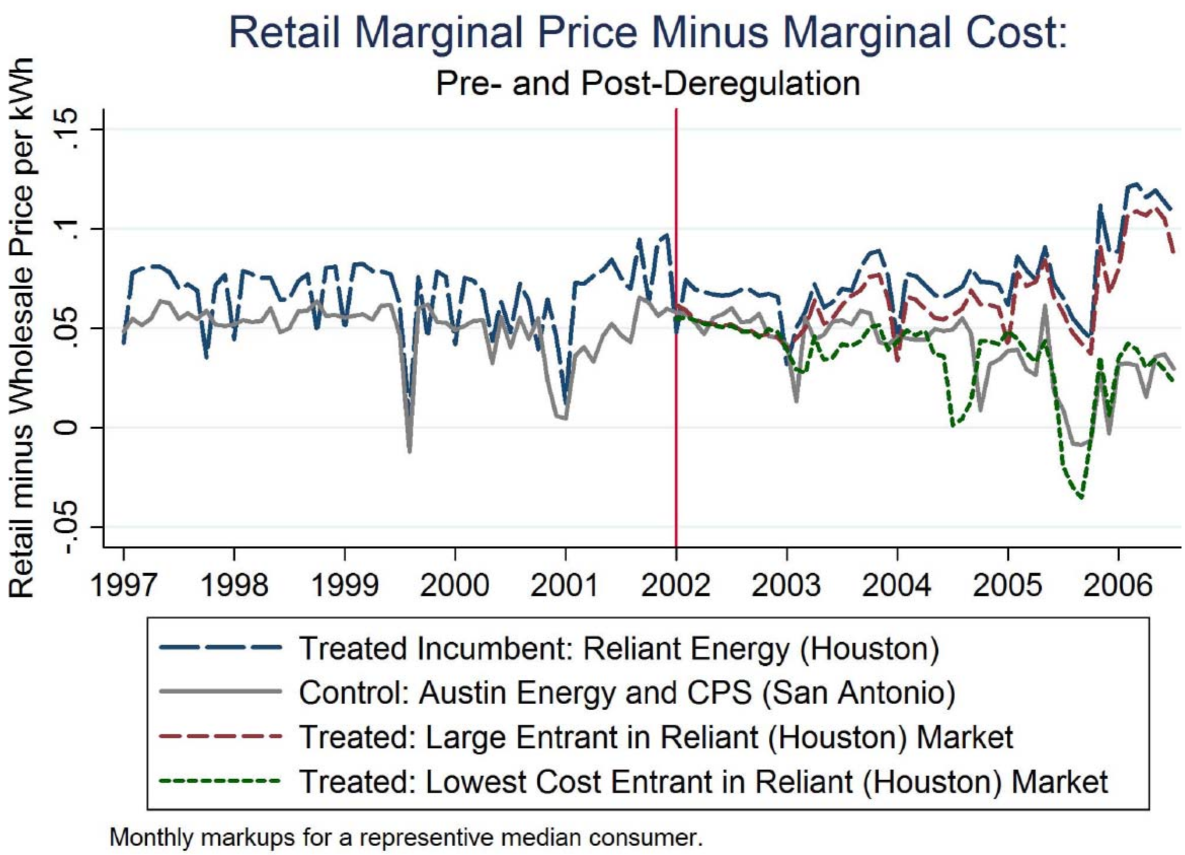
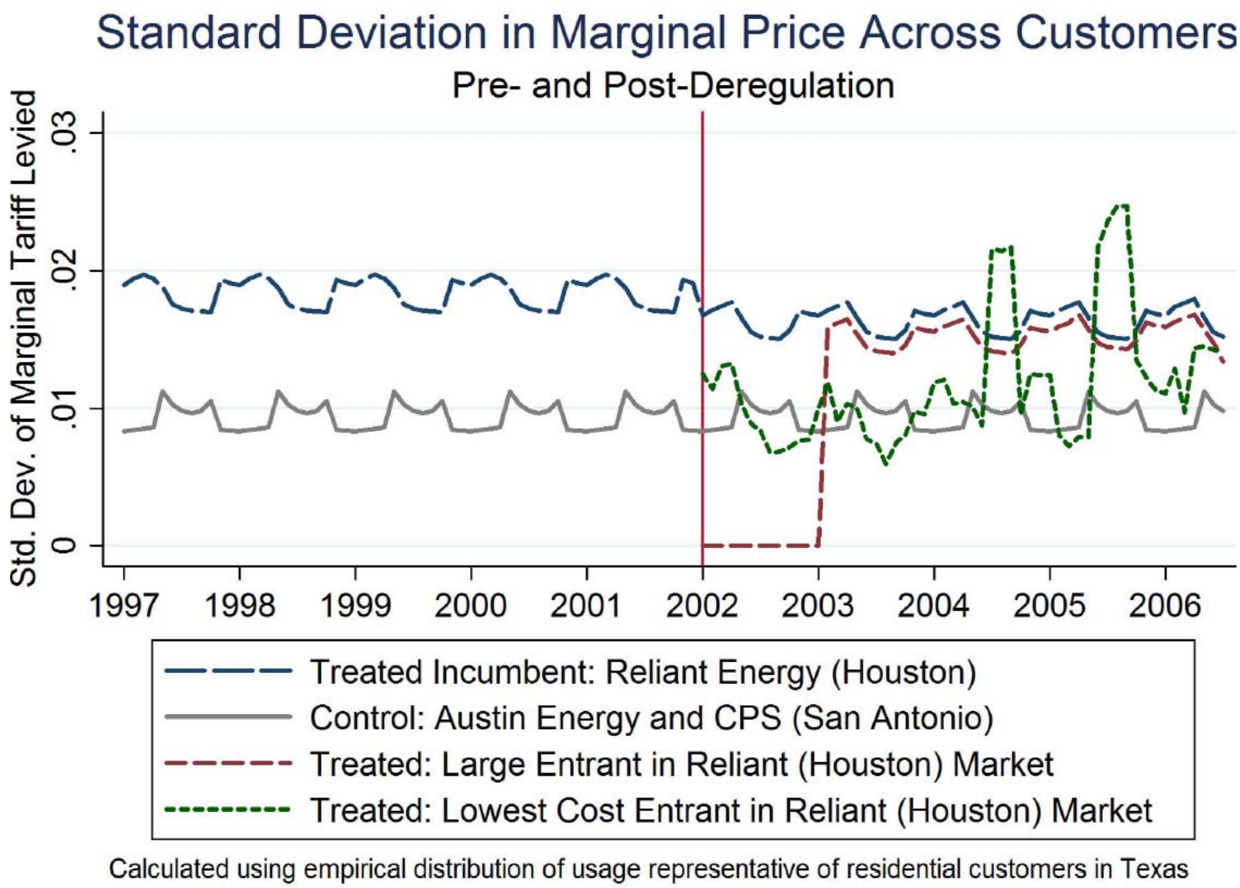

Retail Marginal Price Minus Marginal Cost:
Pre- and Post-Deregulation
Retail minus Wholesale Price per kWh
-.05
0
.05
.1
.15
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
Treated Incumbent: Reliant Energy (Houston)
Control: Austin Energy and CPS (San Antonio)
Treated: Large Entrant in Reliant (Houston) Market
Treated: Lowest Cost Entrant in Reliant (Houston) Market
Monthly markups for a representive median consumer.

Figure 2


Standard Deviation in Marginal Price Across Customers
Pre- and Post-Deregulation
Std. Dev. of Marginal Tariff Levied
0
.01
.02
.03
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
Treated Incumbent: Reliant Energy (Houston)
Control: Austin Energy and CPS (San Antonio)
Treated: Large Entrant in Reliant (Houston) Market
Treated: Lowest Cost Entrant in Reliant (Houston) Market
Calculated using empirical distribution of usage representative of residential customers in Texas