Natural gas elasticities and optimal cost recovery under heterogeneity

Evidence from 300 million natural gas bills in California

Maximilian Auffhammer
UC Berkeley

Edward Rubin
UC Berkeley

Heartland Workshop

September 30, 2017

Natural gas

Gas matters

Natural gas is on the rise.

- Hydraulic fracturing (*fracking*) received significant exemptions from the Clean Air, Clean Water, and Safe Drinking Water Acts via the Energy Policy Act of 2005 (EPA),
- Prices have fallen greatly (*i.e.*, halved) since 2005 (Hausman and Kellogg, 2015),
- Natural gas burns more cleanly and efficiently than other fossil fuels (NAS),
- In 2016, natural gas surpassed coal as the main source of fuel for electricity generation in the U.S. (EIA)—and in CO_2 emissions (EIA).

Natural gas

Gas matters: Consumer edition

Residential natural gas matters too.

- The majority of US households heat primarily with natural gas (AHS),
- Households expend \$50B–\$80B each year on natural gas (EIA & CE),
- Households spend similar amount on their natural gas bills and water bills (CE),
- Natural-gas use is not uniform across income groups (CE).

Motivation

Elasticity of demand

The motivation for this paper stems for two observations:

1. **Policy relevance:** Numerous policy questions require knowledge of the price elasticity of demand for natural gas, *e.g.*,
   - Welfare benefits of natural gas regulation and pricing (*e.g.*, Davis and Muehlegger, 2010)
   - Welfare benefits of fracking (*e.g.*, Hausman and Kellogg, 2015)
2. **Dearth of identified estimates:** Current literature lacks carefully-identified, microdata-based price elasticities of demand for residential natural gas.

Contribution

Overcoming common challenges

- **Two flavors of simultaneity**
  1. Price and quantity result from the equilibrium of a system of equations.

     Problem: Simultaneity bias from failing to separate supply and demand shocks.

     Solutions:

     - Border discontinuity between two utilities
     - Supply instruments: Henry Hub spot price or Eastern US heating degree days
  2. Price is mechanically a function of quantity in multi-tiered pricing regimes.

     Problem: Marginal and average price are endogenous.
     Solution: Proxy/instrument with baseline price or simulated instruments
- **Insufficient data**

  Problem: Lacking consumer-level data on consumption and prices

  Solution: Better data: We combine a large panel of HH bills with utilities' actual prices

Contribution

Overcoming common challenges

- **Two flavors of simultaneity**
  1. Price and quantity result from the equilibrium of a system of equations.

     Problem: Simultaneity bias from failing to separate supply and demand shocks.

     Solutions:
     - Border discontinuity between two utilities
     - Supply instruments: Henry Hub spot price or Eastern US heating degree days
  2. Price is mechanically a function of quantity in multi-tiered pricing regimes.

     Problem: Marginal and average price are endogenous.

     Solution: Proxy/instrument with baseline price or simulated instruments
- **Insufficient data**

  Problem: Lacking consumer-level data on consumption and prices

  Solution: Better data: We combine a large panel of HH bills with utilities' actual prices

Contribution

Overcoming common challenges

- **Two flavors of simultaneity**
  1. Price and quantity result from the equilibrium of a system of equations.

     Problem: Simultaneity bias from failing to separate supply and demand shocks.

     Solutions:

     - Border discontinuity between two utilities
     - Supply instruments: Henry Hub spot price or Eastern US heating degree days
  2. Price is mechanically a function of quantity in multi-tiered pricing regimes.

     Problem: Marginal and average price are endogenous.
     Solution: Proxy/instrument with baseline price or simulated instruments
- **Insufficient data**

  Problem: Lacking consumer-level data on consumption and prices

  Solution: Better data: We combine a large panel of HH bills with utilities' actual prices

Contribution

Overcoming common challenges

- **Two flavors of simultaneity**
  1. Price and quantity result from the equilibrium of a system of equations.

     Problem: Simultaneity bias from failing to separate supply and demand shocks.

     Solutions:

     - Border discontinuity between two utilities
     - Supply instruments: Henry Hub spot price or Eastern US heating degree days
  2. Price is mechanically a function of quantity in multi-tiered pricing regimes.

     Problem: Marginal and average price are endogenous.

     Solution: Proxy/instrument with baseline price or simulated instruments
- **Insufficient data**

  Problem: Lacking consumer-level data on consumption and prices

  Solution: Better data: We combine a large panel of HH bills with utilities' actual prices

Contribution

Combining these empirical strategies with our extensive dataset:

1. We **break simultaneity at the household level**.
2. We are the first paper to **decompose elasticities by season and income**.

This paper

Overview

Main question: What is the elasticity of demand among residential natural gas consumers?

Methods:

- Within-city spatial discontinuities
- Supply-shifting instruments for price

Data: 300M+ natural gas bills from PG&E and SoCalGas

Results:

- We estimate the elast. of demand for residential nat. gas is between -0.29 and -0.21.
- This elasticity varies considerably by season and by income.

Existing literature

Point estimates for the elasticity of demand for residential natural gas

Paper	**Data**	**Estimate**
Houthakker and Taylor (1970)	Time series	–0.15
Herbert and Kreil (1989)	Monthly time series	–0.36
Maddala *et al.* (1997)	US state panel	–0.09 to –0.18
Metcalf and Hassett (1999)	RECS HH panel	–0.08 to –0.71
Garcia–Cerrutti (2000)	Calif. county panel	–0.11
Rehdanz (2007)	Germany HH panel	–0.44 to –0.63
Davis and Muehlegger (2010)	US state panel	–0.278
Meier and Rehdanz (2010)	UK HH panel	–0.34 to –0.56
Hausman and Kellogg (2015)	US state panel	–0.11

Source: Alberini et al. (2011) and authors

Data

Overview

The main datasets in this paper come from 275M+ household bills (all in California).

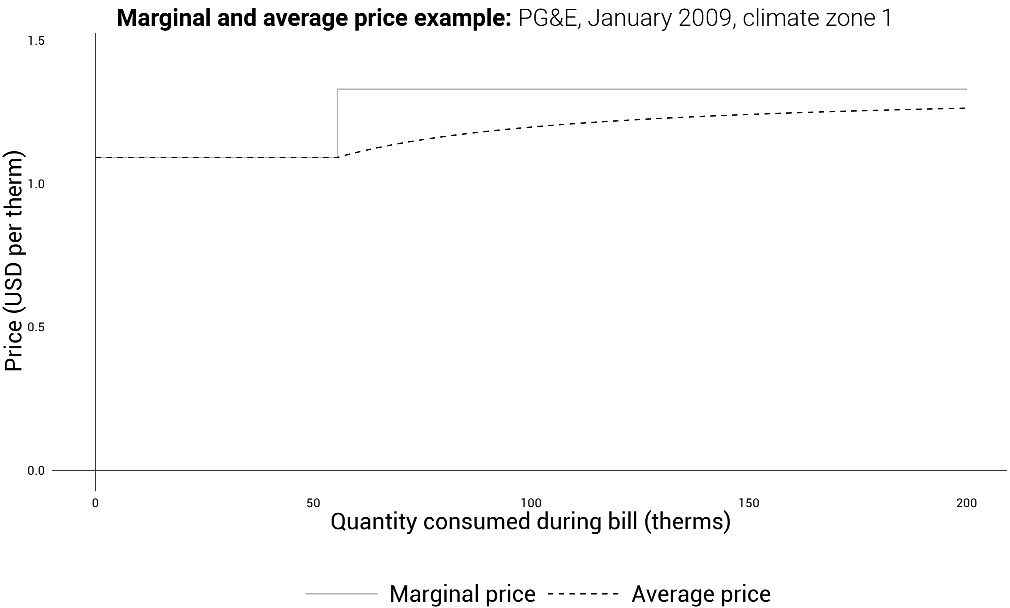
	PG&E	SoCalGas
N. 5-digit zip codes	597	611
N. 9-digit zip codes	680,846	610,207
N. unique households	5,888,276	2,526,503
N. bills	180,663,705	95,335,393
Approx. value (USD)	\$5.71B	\$3.28B

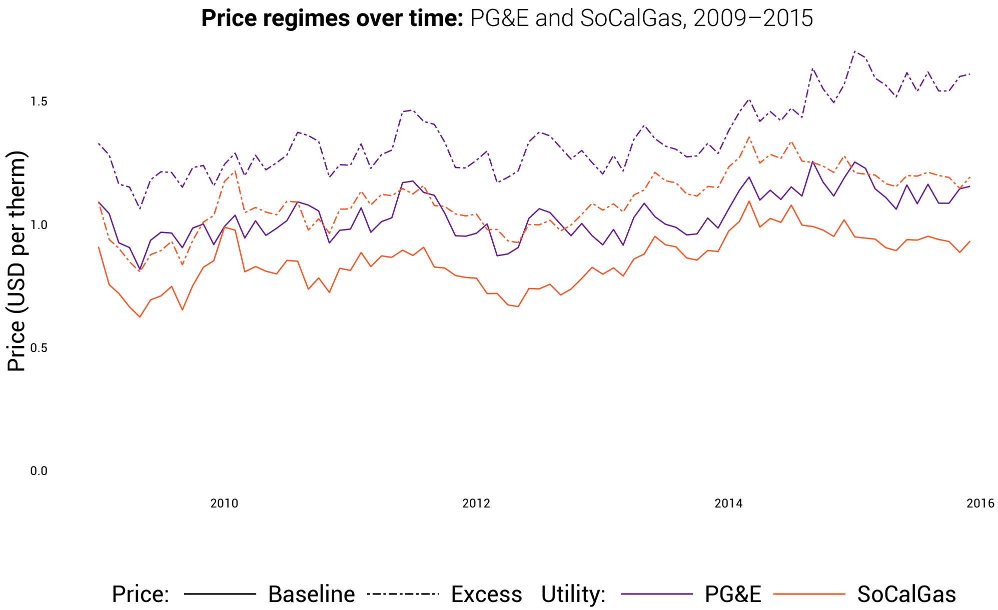
Household information: 9-digit zip, climate zone

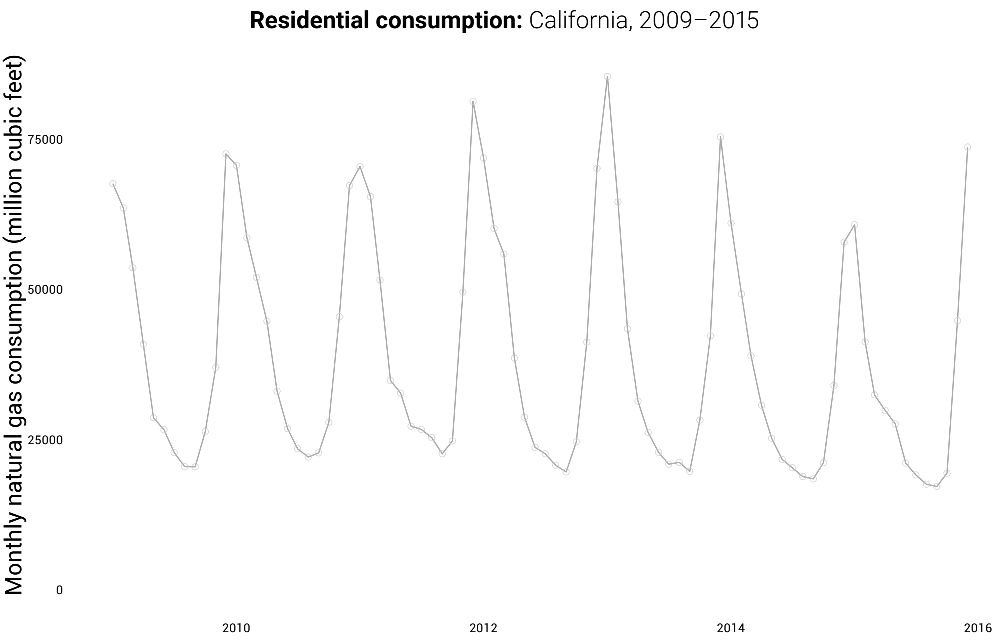
Bill information: dates, quantity, revenue

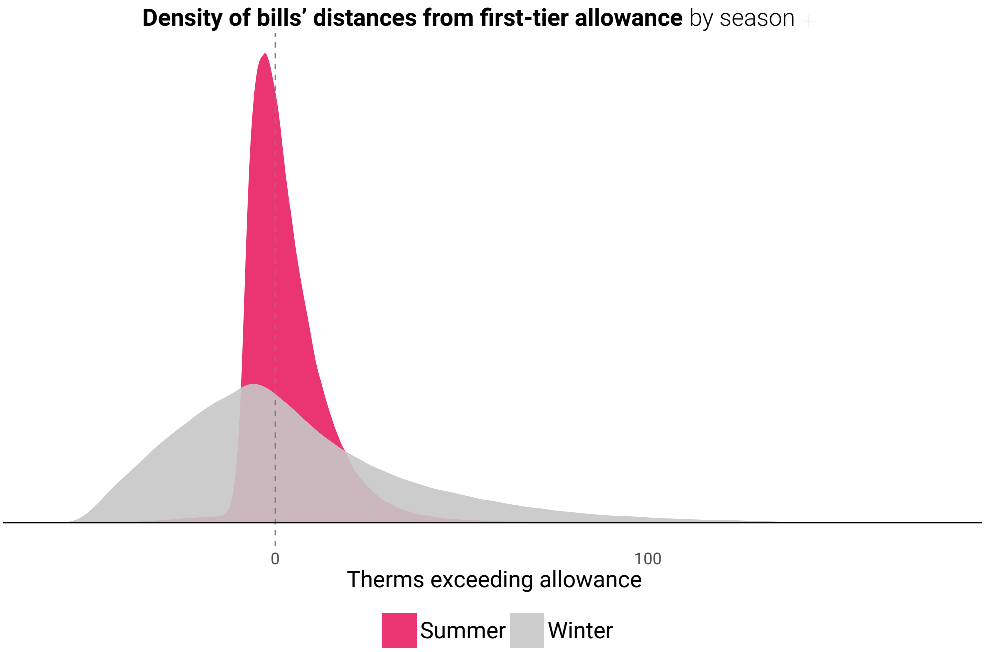
	Overall	Split by season		Split by CARE	
		Winter	Summer	CARE	Non-Care
Baseline price	0.9026	0.8836	0.9204	0.8080	0.9811
Excess price	1.1690	1.1477	1.1891	1.0445	1.2725
Average price	1.0211	1.0008	1.0402	0.9086	1.1147
Marginal price	1.0387	1.0121	1.0637	0.9338	1.1259
Total bill	34.9508	52.0750	18.8573	30.3135	38.8040
Therms	33.8273	50.9544	17.7311	33.1136	34.4204
Exceeded allowance	50.01%	45.93%	53.85%	51.82%	48.51%
Days	30.3994	30.5876	30.2225	30.4040	30.3955
Within-bill HDDs	0.9398	1.7136	0.2126	0.9291	0.9488
(Percent) CARE	45.38%	45.00%	45.74%	100%	0%

Notes: Means of the variables. Summaries based upon the study area. More...


Marginal and average price example: PG&E, January 2009, climate zone 1
Price (USD per therm)
1.5
1.0
0.5
0.0
0
50
100
150
200
Quantity consumed during bill (therms)
Marginal price
Average price


Price regimes over time: PG&E and SoCalGas, 2009–2015
Price (USD per therm)
1.5
1.0
0.5
0.0
2010
2012
2014
2016
Price:
Baseline
Excess
Utility:
PG&E
SoCalGas


Residential consumption: California, 2009–2015
Monthly natural gas consumption (million cubic feet)
75000
50000
25000
0
2010
2012
2014
2016


Density of bills' distances from first-tier allowance by season
0
100
Therms exceeding allowance
Summer
Winter

Empirical strategy

Price elasticity of demand

The relationship at the heart of this paper's elasticity estimates is

$$\log(q_{i,t}) = \eta \log(p_{i,t}) + \gamma_i + \delta_t + \lambda_{i,t} + \varepsilon_{i,t}$$

where *i* and *t* index household and time; *q* denotes quantity consumed; and *p* denotes price.

OLS estimates of η in this equation suffer from two sources of bias/endogeneity:

1. **Simultaneity:** Price and quantity result from the equilibrium of a system of equations.
2. **"Reverse causality":** Two-tiered pricing regime means price is a function of quantity.

Empirical strategy

OLS “elasticity” results

	Dependent variable: Log(Consumption, daily avg.)			
	(1)	(2)	(3)	(4)
Log(Marginal price)	1.0162***	1.0862***		
	(0.0159)	(0.0097)		
Log(Baseline price)			–0.103***	–0.1097***
			(0.0132)	(0.0098)
HH-month FE	T	T	T	T
HDDs	F	T	F	T
Year-month FE	T	F	T	F
City-year-month FE	F	T	F	T
N	16,375,407	16,375,407	16,375,407	16,375,407

Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle.
Significance levels: *10%, **5%, ***1%. Columns represent separate regressions.

Empirical strategy

OLS “elasticity” results

	Dependent variable: Log(Consumption, daily avg.)			
	(1)	(2)	(3)	(4)
Log(Marginal price)	1.0162***	1.0862***		
	(0.0159)	(0.0097)		
Log(Baseline price)			–0.103***	–0.1097***
			(0.0132)	(0.0098)
HH-month FE	T	T	T	T
HDDs	F	T	F	T
Year-month FE	T	F	T	F
City-year-month FE	F	T	F	T
N	16,375,407	16,375,407	16,375,407	16,375,407

Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle.
Significance levels: *10%, **5%, ***1%. Columns represent separate regressions.

Empirical strategy

OLS "elasticity" results

	Dependent variable: Log(Consumption, daily avg.)			
	(1)	(2)	(3)	(4)
Log(Marginal price)	1.0162***	1.0862***		
	(0.0159)	(0.0097)		
Log(Baseline price)			–0.103***	–0.1097***
			(0.0132)	(0.0098)
HH-month FE	T	T	T	T
HDDs	F	T	F	T
Year-month FE	T	F	T	F
City-year-month FE	F	T	F	T
N	16,375,407	16,375,407	16,375,407	16,375,407

Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle.
Significance levels: *10%, **5%, ***1%. Columns represent separate regressions.

Empirical strategy

OLS "elasticity" results

	Dependent variable: Log(Consumption, daily avg.)			
	(1)	(2)	(3)	(4)
Log(Marginal price)	1.0162***	1.0862***		
	(0.0159)	(0.0097)		
Log(Baseline price)			–0.103***	–0.1097***
			(0.0132)	(0.0098)
HH-month FE	T	T	T	T
HDDs	F	T	F	T
Year-month FE	T	F	T	F
City-year-month FE	F	T	F	T
N	16,375,407	16,375,407	16,375,407	16,375,407

Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle.
Significance levels: *10%, **5%, ***1%. Columns represent separate regressions.

Empirical strategy

OLS “elasticity” results

	Dependent variable: Log(Consumption, daily avg.)			
	(1)	(2)	(3)	(4)
Log(Marginal price)	1.0162***	1.0862***		
	(0.0159)	(0.0097)		
Log(Baseline price)			–0.103***	–0.1097***
			(0.0132)	(0.0098)
HH-month FE	T	T	T	T
HDDs	F	T	F	T
Year-month FE	T	F	T	F
City-year-month FE	F	T	F	T
N	16,375,407	16,375,407	16,375,407	16,375,407

Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle.
Significance levels: *10%, **5%, ***1%. Columns represent separate regressions.

Empirical strategy

Potential solutions

1. **Border discontinuity**
2. **Supply-shifting instrument**
3. **Baseline price** or **simulated instruments**

Empirical strategy

Potential solutions

1. **Border discontinuity**
2. **Supply-shifting instrument**
3. **Baseline price** or **simulated instruments**

Empirical strategy

Border discontinuity

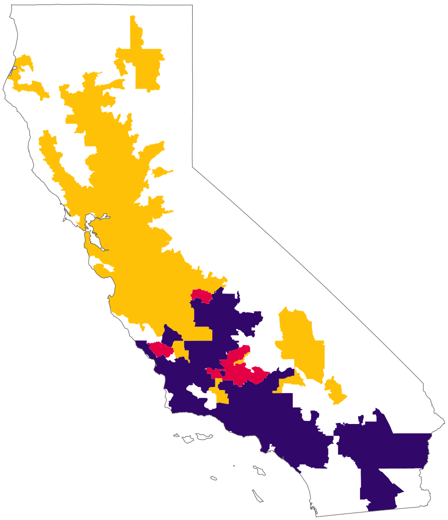
Discontinuity: The border between PG&E and SoCalGas bisects 11 cities (39 zip codes) in southern California (akin to Ito, 2014).

Motivation: Represents edge of long-established underground networks of pipes.

Identification: If the boundary is orthogonal to consumers' preferences, then we can identify off of the within-city differences in prices and consumption across this discontinuity (*i.e.*, one side controls for the other).

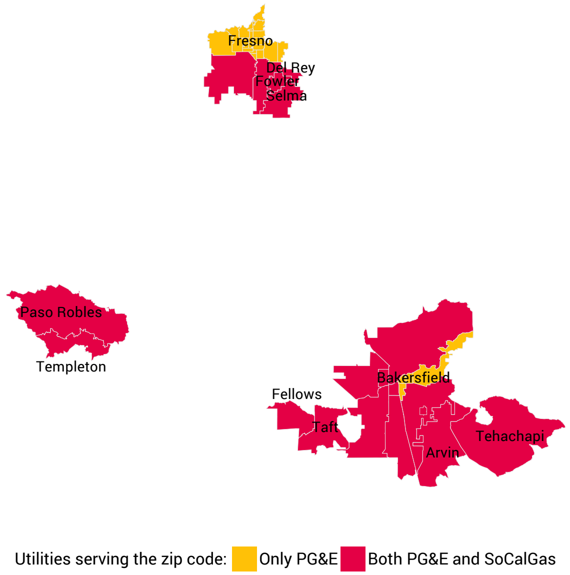
Concern: Sorting.

Natural gas service areas: data coverage by zip code



Utility presence: PG&E SoCalGas PG&E and SoCalGas

Study-area discontinuity: Zip codes in cities served by both utilities



Supply-shifting instrument

Henry Hub spot-price instrument

Instrument: Avg. spot price at Louisiana's Henry Hub in the week preceding price changes, interacted with utility.

Motivation: Reflects national price of natural gas: utilities purchase gas from a nationally integrated market. Utilities are rate-of-return earners.

Identification: Valid instrument if

1. Spot prices predict residential prices, and
2. Spot prices are uncorrelated with demand shocks, after controlling for within-bill HDDs and zip-code by month-of-sample fixed effects.

Concern: CA utilities could have market power due to other gas uses.

Empirical strategy

Henry Hub spot-price instrument

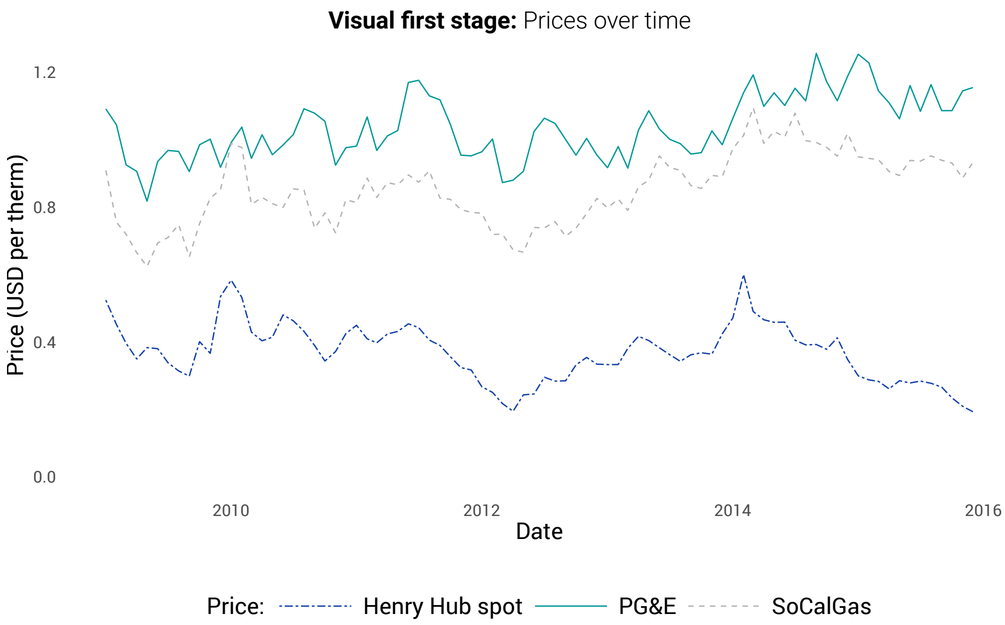
The first- and second-stages for this IV strategy:

$$\log\left(p_{i,t}\right) = \pi_{1a} p_{i,t}^{\text{spot}} + \pi_{1b} p_{i,t}^{\text{spot}} \times \text{SCG}_i + \pi_2 \text{HDD}_{i,t}^{\text{bill}} + \text{HH}_{i,t} + \text{Zip}_{i,t} + u_{i,t}$$

$$\log\left(q_{i,t}\right) = \eta_1 \widehat{\log\left(p_{i,t}\right)} + \eta_2 \text{HDD}_{i,t}^{\text{bill}} + \text{HH}_{i,t} + \text{Zip}_{i,t} + v_{i,t}$$

where

$p_{i,t}$	household i's price in period t
$p_{i,t}^{\text{spot}}$	the spot price in the weeks preceding i's utility setting $p_{i,t}$
SCG_i	indicator for whether household i's utility is SoCalGas
$\text{HDD}_{i,t}^{\text{bill}}$	number of HDDs during household i's bill in time period t
$\text{HH}_{i,t}$	fixed effect for household i in month-of-year t
$\text{Zip}_{i,t}$	fixed effect for i's zip in month-of-sample t
$q_{i,t}$	household i's average daily consumption during their bill in t


Visual first stage: Prices over time
Price (USD per therm)
1.2
0.8
0.4
0.0
2010
2012
2014
2016
Date
Price:
Henry Hub spot
PG&E
SoCalGas

First-stage results

Henry Hub spot price IV

	Dependent variable (endogenous price) in first stage:					
	Log(Marginal price)		Log(Average price)		Log(Baseline price)	
Spot price	0.2174*	0.2391**	0.3299***	0.3416***	0.5345***	0.5383***
	(0.1268)	(0.113)	(0.0956)	(0.089)	(0.0669)	(0.0653)
Spot price × SoCalGas	0.7835***	0.7816***	0.7217***	0.7207***	0.8361***	0.8358***
	(0.0275)	(0.0282)	(0.0236)	(0.0238)	(0.0232)	(0.0232)
F stat. for instruments	482.7	442.7	558.1	541.6	788.9	791.7
Bill's HDD	F	T	F	T	F	T
HH-month FE	T	T	T	T	T	T
Zip-year-month FE	T	T	T	T	T	T
N	16,375,407	16,375,407	16,375,407	16,375,407	16,375,407	16,375,407

Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle.
Significance levels: *10%, **5%, ***1%. Columns represent separate regressions.

First-stage results

Henry Hub spot price IV

	Dependent variable (endogenous price) in first stage:					
	Log(Marginal price)		Log(Average price)		Log(Baseline price)	
Spot price	0.2174*	0.2391**	0.3299***	0.3416***	0.5345***	0.5383***
	(0.1268)	(0.113)	(0.0956)	(0.089)	(0.0669)	(0.0653)
Spot price × SoCalGas	0.7835***	0.7816***	0.7217***	0.7207***	0.8361***	0.8358***
	(0.0275)	(0.0282)	(0.0236)	(0.0238)	(0.0232)	(0.0232)
F stat. for instruments	482.7	442.7	558.1	541.6	788.9	791.7
Bill's HDD	F	T	F	T	F	T
HH-month FE	T	T	T	T	T	T
Zip-year-month FE	T	T	T	T	T	T
N	16,375,407	16,375,407	16,375,407	16,375,407	16,375,407	16,375,407

Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle.
Significance levels: *10%, **5%, ***1%. Columns represent separate regressions.

First-stage results

Henry Hub spot price IV

	Dependent variable (endogenous price) in first stage:					
	Log(Marginal price)		Log(Average price)		Log(Baseline price)	
Spot price	0.2174*	0.2391**	0.3299***	0.3416***	0.5345***	0.5383***
	(0.1268)	(0.113)	(0.0956)	(0.089)	(0.0669)	(0.0653)
Spot price × SoCalGas	0.7835***	0.7816***	0.7217***	0.7207***	0.8361***	0.8358***
	(0.0275)	(0.0282)	(0.0236)	(0.0238)	(0.0232)	(0.0232)
F stat. for instruments	482.7	442.7	558.1	541.6	788.9	791.7
Bill's HDD	F	T	F	T	F	T
HH-month FE	T	T	T	T	T	T
Zip-year-month FE	T	T	T	T	T	T
N	16,375,407	16,375,407	16,375,407	16,375,407	16,375,407	16,375,407

Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle.
Significance levels: *10%, **5%, ***1%. Columns represent separate regressions.

First-stage results

Henry Hub spot price IV

	Dependent variable (endogenous price) in first stage:					
	Log(Marginal price)		Log(Average price)		Log(Baseline price)	
Spot price	0.2174*	0.2391**	0.3299***	0.3416***	0.5345***	0.5383***
	(0.1268)	(0.113)	(0.0956)	(0.089)	(0.0669)	(0.0653)
Spot price × SoCalGas	0.7835***	0.7816***	0.7217***	0.7207***	0.8361***	0.8358***
	(0.0275)	(0.0282)	(0.0236)	(0.0238)	(0.0232)	(0.0232)
F stat. for instruments	482.7	442.7	558.1	541.6	788.9	791.7
Bill's HDD	F	T	F	T	F	T
HH-month FE	T	T	T	T	T	T
Zip-year-month FE	T	T	T	T	T	T
N	16,375,407	16,375,407	16,375,407	16,375,407	16,375,407	16,375,407

Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle.
Significance levels: *10%, **5%, ***1%. Columns represent separate regressions.

First-stage results

Henry Hub spot price IV

	Dependent variable (endogenous price) in first stage:					
	Log(Marginal price)		Log(Average price)		Log(Baseline price)	
Spot price	0.2174*	0.2391**	0.3299***	0.3416***	0.5345***	0.5383***
	(0.1268)	(0.113)	(0.0956)	(0.089)	(0.0669)	(0.0653)
Spot price × SoCalGas	0.7835***	0.7816***	0.7217***	0.7207***	0.8361***	0.8358***
	(0.0275)	(0.0282)	(0.0236)	(0.0238)	(0.0232)	(0.0232)
F stat. for instruments	482.7	442.7	558.1	541.6	788.9	791.7
Bill's HDD	F	T	F	T	F	T
HH-month FE	T	T	T	T	T	T
Zip-year-month FE	T	T	T	T	T	T
N	16,375,407	16,375,407	16,375,407	16,375,407	16,375,407	16,375,407

Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle.
Significance levels: *10%, **5%, ***1%. Columns represent separate regressions.

Elasticity results

Second-stage results

First Stage Henry Hub spot price instrument

Dependent variable: Log(Consumption, daily avg.)					
	Type of price: instrumented explanatory variable				
	Mrg.	Sim. Mrg.	Avg. Mrg.	Avg.	Baseline
Log(Price)	–0.2125***	–0.2581***	–0.2464***	–0.2902***	–0.2903***
instrumented	(0.0611)	(0.0622)	(0.0563)	(0.0689)	(0.0608)
Within-bill HDDs	0.3617***	0.3686***	0.3596***	0.3651***	0.3716***
(thousands)	(0.0153)	(0.0159)	(0.0149)	(0.0152)	(0.0152)
F stat. for instruments	442.7	525.1	722.6	541.6	791.7
Bill's HDD	T	T	T	T	T
HH-month FE	T	T	T	T	T
Zip-year-month FE	T	T	T	T	T
N	16,375,407	13,675,986	16,375,407	16,375,407	16,375,407

Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle.
Significance levels: *10%, **5%, ***1%. Columns represent separate regressions.

Second-stage results

First Stage Henry Hub spot price instrument

Dependent variable: Log(Consumption, daily avg.)					
	Type of price: instrumented explanatory variable				
	Mrg.	Sim. Mrg.	Avg. Mrg.	Avg.	Baseline
Log(Price)	-0.2125^{***}	-0.2581^{***}	-0.2464^{***}	-0.2902^{***}	-0.2903^{***}
instrumented	(0.0611)	(0.0622)	(0.0563)	(0.0689)	(0.0608)
Within-bill HDDs	0.3617^{***}	0.3686^{***}	0.3596^{***}	0.3651^{***}	0.3716^{***}
(thousands)	(0.0153)	(0.0159)	(0.0149)	(0.0152)	(0.0152)
F stat. for instruments	442.7	525.1	722.6	541.6	791.7
Bill's HDD	T	T	T	T	T
HH-month FE	T	T	T	T	T
Zip-year-month FE	T	T	T	T	T
N	16,375,407	13,675,986	16,375,407	16,375,407	16,375,407

Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle.
Significance levels: *10%, **5%, ***1%. Columns represent separate regressions.

Second-stage results

First Stage Henry Hub spot price instrument

	Dependent variable: Log(Consumption, daily avg.)				
	Type of price: instrumented explanatory variable				
	Mrg.	Sim. Mrg.	Avg. Mrg.	Avg.	Baseline
Log(Price)	-0.2125^{***}	-0.2581^{***}	-0.2464^{***}	-0.2902^{***}	-0.2903^{***}
instrumented	(0.0611)	(0.0622)	(0.0563)	(0.0689)	(0.0608)
Within-bill HDDs	0.3617^{***}	0.3686^{***}	0.3596^{***}	0.3651^{***}	0.3716^{***}
(thousands)	(0.0153)	(0.0159)	(0.0149)	(0.0152)	(0.0152)
F stat. for instruments	442.7	525.1	722.6	541.6	791.7
Bill's HDD	T	T	T	T	T
HH-month FE	T	T	T	T	T
Zip-year-month FE	T	T	T	T	T
N	16,375,407	13,675,986	16,375,407	16,375,407	16,375,407

Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle.
Significance levels: *10%, **5%, ***1%. Columns represent separate regressions.

Second-stage results

First Stage Henry Hub spot price instrument

Dependent variable: Log(Consumption, daily avg.)

	Type of price: instrumented explanatory variable				
	Mrg.	Sim. Mrg.	Avg. Mrg.	Avg.	Baseline
Log(Price)	-0.2125^{***}	-0.2581^{***}	-0.2464^{***}	-0.2902^{***}	-0.2903^{***}
instrumented	(0.0611)	(0.0622)	(0.0563)	(0.0689)	(0.0608)
Within-bill HDDs	0.3617^{***}	0.3686^{***}	0.3596^{***}	0.3651^{***}	0.3716^{***}
(thousands)	(0.0153)	(0.0159)	(0.0149)	(0.0152)	(0.0152)
F stat. for instruments	442.7	525.1	722.6	541.6	791.7
Bill's HDD	T	T	T	T	T
HH-month FE	T	T	T	T	T
Zip-year-month FE	T	T	T	T	T
N	16,375,407	13,675,986	16,375,407	16,375,407	16,375,407

Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle.
Significance levels: *10%, **5%, ***1%. Columns represent separate regressions.

What about heterogeneity?

Income heterogeneity

2SLS results for Cal. Alt. Rates for Energy (CARE) status

Dependent variable: Log(Consumption, daily avg.)

	Income group			
	CARE households		Non-CARE households	
Log(Marginal price)	−0.2276***	−0.2656***	−0.1255**	−0.1372***
instrumented	(0.0716)	(0.0570)	(0.0585)	(0.0483)
Within-bill HDDs		0.3112***		0.3493***
(thousands)		(0.0173)		(0.0184)
F stat. for instruments	727.1	723.6	795.3	750.1
Zip code's HDDs during bill	F	T	F	T
HH-month FE	T	T	T	T
City-year-month FE	T	T	T	T
N	7,431,500	7,431,500	8,943,907	8,943,907

Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle.
Significance levels: *10%, **5%, ***1%. Columns represent separate regressions.

Income heterogeneity

2SLS results for Cal. Alt. Rates for Energy (CARE) status

	Dependent variable: Log(Consumption, daily avg.)			
	Income group			
	CARE households		Non-CARE households	
Log(Marginal price)	-0.2276^{***}	-0.2656^{***}	-0.1255^{**}	-0.1372^{***}
instrumented	(0.0716)	(0.0570)	(0.0585)	(0.0483)
Within-bill HDDs		0.3112^{***}		0.3493^{***}
(thousands)		(0.0173)		(0.0184)
F stat. for instruments	727.1	723.6	795.3	750.1
Zip code's HDDs during bill	F	T	F	T
HH-month FE	T	T	T	T
City-year-month FE	T	T	T	T
N	7,431,500	7,431,500	8,943,907	8,943,907

Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle.
Significance levels: *10%, **5%, ***1%. Columns represent separate regressions.

Income heterogeneity

2SLS results for Cal. Alt. Rates for Energy (CARE) status

Dependent variable: Log(Consumption, daily avg.)				
	Income group			
	CARE households		Non-CARE households	
Log(Marginal price) *instrumented*	−0.2276*** (0.0716)	$\boxed{-0.2656^{***}\ (0.0570)}$	−0.1255** (0.0585)	$\boxed{-0.1372^{***}\ (0.0483)}$
Within-bill HDDs (thousands)		0.3112*** (0.0173)		0.3493*** (0.0184)
F stat. for instruments	727.1	723.6	795.3	750.1
Zip code's HDDs during bill	F	T	F	T
HH-month FE	T	T	T	T
City-year-month FE	T	T	T	T
N	7,431,500	7,431,500	8,943,907	8,943,907

Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle.
Significance levels: *10%, **5%, ***1%. Columns represent separate regressions.

Income heterogeneity

2SLS results for Cal. Alt. Rates for Energy (CARE) status

Dependent variable: Log(Consumption, daily avg.)				
	Income group			
	CARE households		Non-CARE households	
Log(Marginal price)	-0.2276^{***}	-0.2656^{***}	-0.1255^{**}	-0.1372^{***}
instrumented	(0.0716)	(0.0570)	(0.0585)	(0.0483)
Within-bill HDDs		0.3112^{***}		0.3493^{***}
(thousands)		(0.0173)		(0.0184)
F stat. for instruments	727.1	723.6	795.3	750.1
Zip code's HDDs during bill	F	T	F	T
HH-month FE	T	T	T	T
City-year-month FE	T	T	T	T
N	7,431,500	7,431,500	8,943,907	8,943,907

Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle.
Significance levels: *10%, **5%, ***1%. Columns represent separate regressions.

Seasonal heterogeneity

2SLS results

Dependent variable: Log(Consumption, daily avg.)

	Season			
	Winter months		Summer months	
Log(Marginal price)	-0.4153^{***}	-0.4736^{***}	-0.0384	-0.0382
instrumented	(0.1159)	(0.1064)	(0.0478)	(0.0336)
Within-bill HDDs		0.3201^{***}		0.2835^{***}
(thousands)		(0.0216)		(0.0157)
F stat. for instruments	265.1	252.3	922.1	929.6
Zip code's HDDs during bill	F	T	F	T
HH-month FE	T	T	T	T
City-year-month FE	T	T	T	T
N	7,933,660	7,933,660	8,441,747	8,441,747

Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle.
Significance levels: *10%, **5%, ***1%. Columns represent separate regressions.

Seasonal heterogeneity

2SLS results

Dependent variable: Log(Consumption, daily avg.)				
	Season			
	Winter months		Summer months	
Log(Marginal price)	-0.4153^{***}	-0.4736^{***}	-0.0384	-0.0382
instrumented	(0.1159)	(0.1064)	(0.0478)	(0.0336)
Within-bill HDDs		0.3201^{***}		0.2835^{***}
(thousands)		(0.0216)		(0.0157)
F stat. for instruments	265.1	252.3	922.1	929.6
Zip code's HDDs during bill	F	T	F	T
HH-month FE	T	T	T	T
City-year-month FE	T	T	T	T
N	7,933,660	7,933,660	8,441,747	8,441,747

Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle.
Significance levels: *10%, **5%, ***1%. Columns represent separate regressions.

Seasonal heterogeneity

2SLS results

Dependent variable: Log(Consumption, daily avg.)

	Season			
	Winter months		Summer months	
Log(Marginal price) *instrumented*	-0.4153^{***} (0.1159)	-0.4736^{***} (0.1064)	-0.0384 (0.0478)	-0.0382 (0.0336)
Within-bill HDDs (thousands)		0.3201^{***} (0.0216)		0.2835^{***} (0.0157)
F stat. for instruments	265.1	252.3	922.1	929.6
Zip code's HDDs during bill	F	T	F	T
HH-month FE	T	T	T	T
City-year-month FE	T	T	T	T
N	7,933,660	7,933,660	8,441,747	8,441,747

Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle.
Significance levels: *10%, **5%, ***1%. Columns represent separate regressions.

Heterogeneity by income and season

2SLS results

Dependent variable: Log(Consumption, daily avg.)	Summer CARE	Summer Non-CARE	Winter CARE	Winter Non-CARE
Log(Marginal price)	–0.0797*	–0.0208	–0.6062***	–0.3402***
instrumented	(0.0407)	(0.0412)	(0.1343)	(0.1133)
Within-bill HDDs	0.2652***	0.3014***	0.2927***	0.3519***
(thousands)	(0.0195)	(0.0165)	(0.0233)	(0.0235)
F stat. for instruments	737.5	787	196.4	248.8
Zip code's HDDs during bill	T	T	T	T
HH-month FE	T	T	T	T
City-year-month FE	T	T	T	T
N	3,861,459	4,580,288	3,570,041	4,363,619

Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle.
Significance levels: *10%, **5%, ***1%. Columns represent separate regressions.

Heterogeneity by income and season

2SLS results

Dependent variable: Log(Consumption, daily avg.)

	Summer CARE	Summer Non-CARE	Winter CARE	Winter Non-CARE
Log(Marginal price)	−0.0797*	−0.0208	−0.6062***	−0.3402***
instrumented	(0.0407)	(0.0412)	(0.1343)	(0.1133)
Within-bill HDDs	0.2652***	0.3014***	0.2927***	0.3519***
(thousands)	(0.0195)	(0.0165)	(0.0233)	(0.0235)
F stat. for instruments	737.5	787	196.4	248.8
Zip code's HDDs during bill	T	T	T	T
HH-month FE	T	T	T	T
City-year-month FE	T	T	T	T
N	3,861,459	4,580,288	3,570,041	4,363,619

Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle.
Significance levels: *10%, **5%, ***1%. Columns represent separate regressions.

Heterogeneity by income and season

2SLS results

Dependent variable: Log(Consumption, daily avg.)

	Summer CARE	Summer Non-CARE	Winter CARE	Winter Non-CARE
Log(Marginal price)	−0.0797*	−0.0208	−0.6062***	−0.3402***
instrumented	(0.0407)	(0.0412)	(0.1343)	(0.1133)
Within-bill HDDs	0.2652***	0.3014***	0.2927***	0.3519***
(thousands)	(0.0195)	(0.0165)	(0.0233)	(0.0235)
F stat. for instruments	737.5	787	196.4	248.8
Zip code's HDDs during bill	T	T	T	T
HH-month FE	T	T	T	T
City-year-month FE	T	T	T	T
N	3,861,459	4,580,288	3,570,041	4,363,619

Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle.
Significance levels: *10%, **5%, ***1%. Columns represent separate regressions.

Heterogeneity by income and season

2SLS results

Dependent variable: Log(Consumption, daily avg.)

	Summer CARE	Summer Non-CARE	Winter CARE	Winter Non-CARE
Log(Marginal price)	–0.0797*	–0.0208	–0.6062***	–0.3402***
instrumented	(0.0407)	(0.0412)	(0.1343)	(0.1133)
Within-bill HDDs	0.2652***	0.3014***	0.2927***	0.3519***
(thousands)	(0.0195)	(0.0165)	(0.0233)	(0.0235)
F stat. for instruments	737.5	787	196.4	248.8
Zip code's HDDs during bill	T	T	T	T
HH-month FE	T	T	T	T
City-year-month FE	T	T	T	T
N	3,861,459	4,580,288	3,570,041	4,363,619

Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle.
Significance levels: *10%, **5%, ***1%. Columns represent separate regressions.

Heterogeneity by income and season

2SLS results

Dependent variable: Log(Consumption, daily avg.)

	Summer CARE	Summer Non-CARE	Winter CARE	Winter Non-CARE
Log(Marginal price)	-0.0797^{*}	-0.0208	-0.6062^{***}	-0.3402^{***}
instrumented	(0.0407)	(0.0412)	(0.1343)	(0.1133)
Within-bill HDDs	0.2652^{***}	0.3014^{***}	0.2927^{***}	0.3519^{***}
(thousands)	(0.0195)	(0.0165)	(0.0233)	(0.0235)
F stat. for instruments	737.5	787	196.4	248.8
Zip code's HDDs during bill	T	T	T	T
HH-month FE	T	T	T	T
City-year-month FE	T	T	T	T
N	3,861,459	4,580,288	3,570,041	4,363,619

Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle.
Significance levels: *10%, **5%, ***1%. Columns represent separate regressions.

Heterogeneity by income and season

2SLS results

Dependent variable: Log(Consumption, daily avg.)

	Summer CARE	Summer Non-CARE	Winter CARE	Winter Non-CARE
Log(Marginal price)	-0.0797^{*}	-0.0208	-0.6062^{***}	-0.3402^{***}
instrumented	(0.0407)	(0.0412)	(0.1343)	(0.1133)
Within-bill HDDs	0.2652^{***}	0.3014^{***}	0.2927^{***}	0.3519^{***}
(thousands)	(0.0195)	(0.0165)	(0.0233)	(0.0235)
F stat. for instruments	737.5	787	196.4	248.8
Zip code's HDDs during bill	T	T	T	T
HH-month FE	T	T	T	T
City-year-month FE	T	T	T	T
N	3,861,459	4,580,288	3,570,041	4,363,619

Errors are two-way clustered within (1) household and (2) utility by climate-zone by billing-cycle.
Significance levels: *10%, **5%, ***1%. Columns represent separate regressions.

Discussion

Utilizing elasticity heterogeneity

Motivating observation: Utilities, public utility commissions, and local governments add fees and taxes to bills throughout the year.

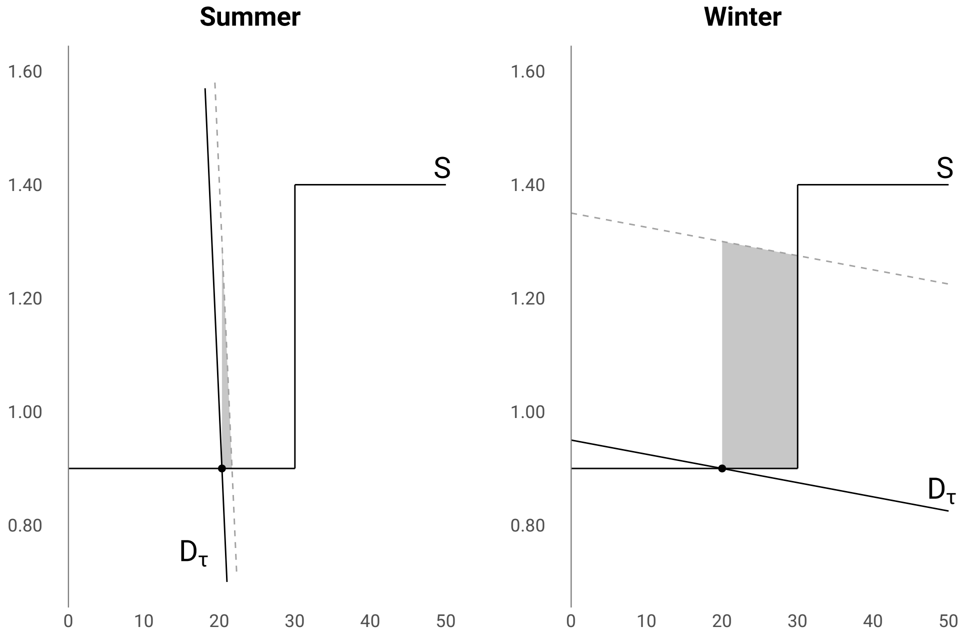
In California, most of these fees and taxes are volumetric.

Discussion

Policy implication: By shifting fees and taxes to the less elastic times of the year (*i.e.*, the summer), utilities and governments can decrease the dead-weight losses associated with these charges (**increased efficiency**).

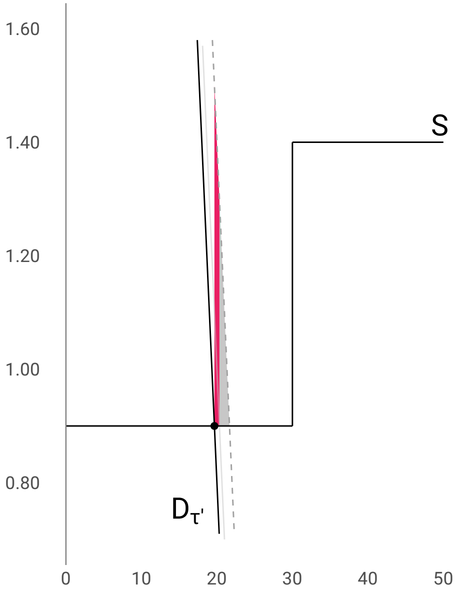
Further, because CARE households are substantially more elastic in the winter, this shift is **potentially progressive**.

While this suggestion is particularly relevant for utilities/governments that recover costs through volumetric charges (*e.g.*, California), if consumers respond to average prices or total bill, then this suggestion is relevant for *all* fees that consumers face.

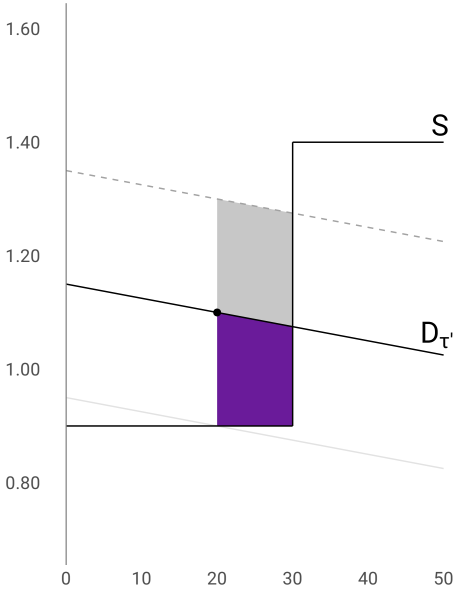


Summer
Winter
S
S
D_τ
D_τ
1.60
1.40
1.20
1.00
0.80
0
10
20
30
40
50

Summer: ↑ tax by $0.20 ⇒ ↑ DWL



Winter: ↓ tax by $0.20 ⇒ ↓ DWL



Conclusion

“Pooled” results: Through a variety of specifications, our point estimates for the “pooled” price elasticity of demand for residential natural gas range from –0.29 to –0.21.

Heterogeneity: However, we find significant evidence of heterogeneity within this elasticity—both with respect to season and with respect to income.

Implications: Taking this heterogeneity into account offers

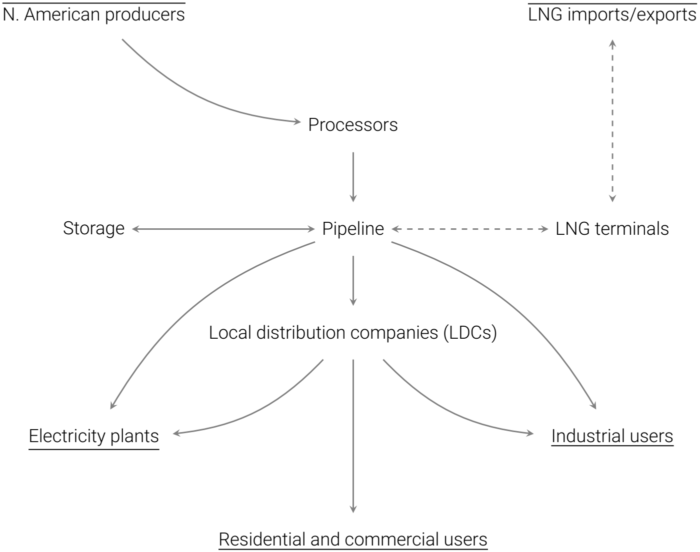
1. Empirical insights into heterogeneity underlying more standard “pooled” elasticities, and
2. Unexplored avenues for potentially more efficient and progressive policies.

Thank you!

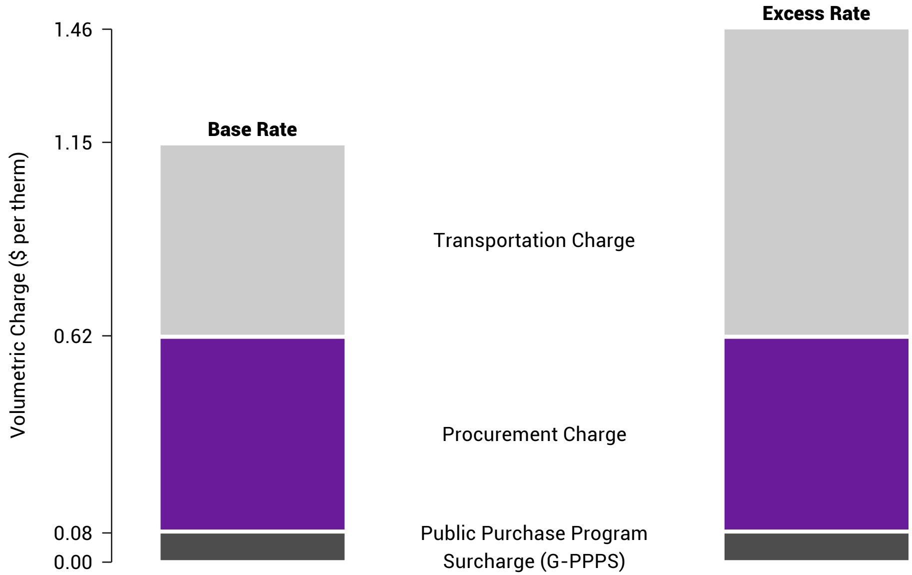
Ed Rubin
edward@berkeley.edu
http://edrub.in

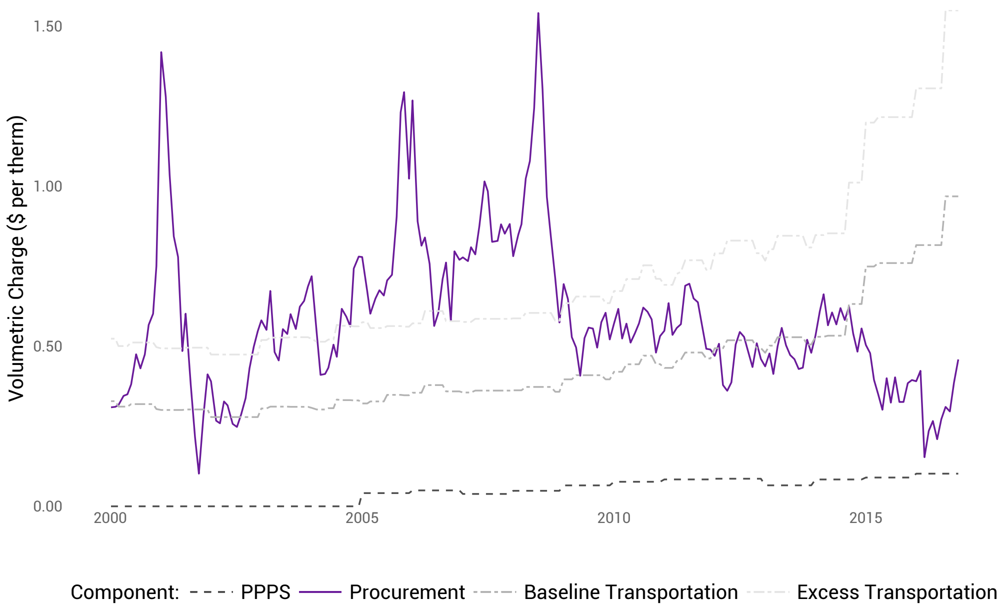
Institutional context

Natural gas system



Note: We based this figure off Levine, Carpenter, and Thapa (2014), with some modification following Brown and Yücel (1993).


Excess Rate
Base Rate
1.46
1.15
0.62
0.08
0.00
Volumetric Charge ($ per therm)
Transportation Charge
Procurement Charge
Public Purchase Program
Surcharge (G-PPPS)



Volumetric Charge ($ per therm)
1.50
1.00
0.50
0.00
2000
2005
2010
2015
Component:
PPPS
Procurement
Baseline Transportation
Excess Transportation

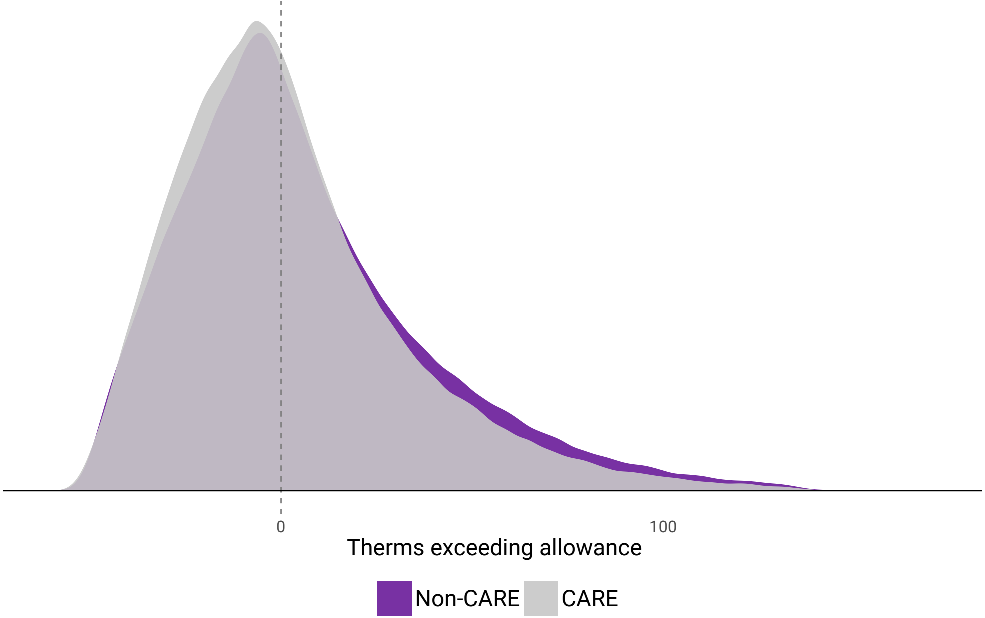
		Split by season		Split by CARE	
	Overall	Winter	Summer	CARE	Non-Care
Baseline price	0.9026	0.8836	0.9204	0.8080	0.9811
	[0.1419]	[0.1361]	[0.1448]	[0.0854]	[0.1311]
Excess price	1.1690	1.1477	1.1891	1.0445	1.2725
	[0.1742]	[0.1708]	[0.1751]	[0.1009]	[0.1534]
Average price	1.0211	1.0008	1.0402	0.9086	1.1147
	[0.1621]	[0.1583]	[0.1633]	[0.1004]	[0.1430]
Marginal price	1.0387	1.0121	1.0637	0.9338	1.1259
	[0.1983]	[0.1905]	[0.2021]	[0.1448]	[0.1944]

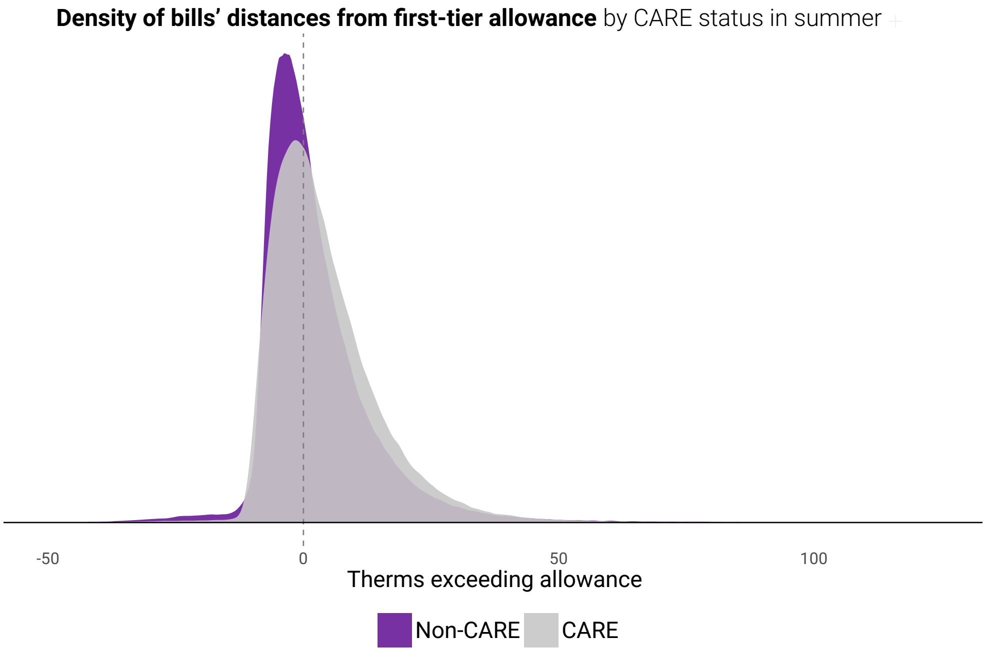
Notes: Unbracketed values provide the means of the variables; bracketed values denote the variables' standard deviations. Summaries based upon the study area. Back.

		Split by season		Split by CARE	
	Overall	Winter	Summer	CARE	Non-Care
Therms	33.8273	50.9544	17.7311	33.1136	34.4204
	[30.7697]	[35.2487]	[11.5803]	[28.7629]	[32.3306]
Days	30.3994	30.5876	30.2225	30.4040	30.3955
	[1.3038]	[1.3843]	[1.1966]	[1.2761]	[1.3263]
Therms per day	1.1063	1.6588	0.5871	1.0840	1.1249
	[0.9936]	[1.1354]	[0.3838]	[0.9304]	[1.0429]
Total bill	34.9508	52.0750	18.8573	30.3135	38.8040
	[33.8812]	[39.8973]	[14.0069]	[27.2567]	[38.1017]
Within-bill HDDs	0.9398	1.7136	0.2126	0.9291	0.9488
	[0.8445]	[0.4219]	[0.3529]	[0.8443]	[0.8446]
(Percent) CARE	45.38%	45.00%	45.74%	100%	0%

Notes: Unbracketed values provide the means of the variables; bracketed values denote the variables' standard deviations. Summaries based upon the study area. Back.

Density of bills' distances from first-tier allowance by CARE status in winter





Density of bills' distances from first-tier allowance by CARE status in summer
-50
0
50
100
Therms exceeding allowance
Non-CARE
CARE

Definition: Instrument current price with value(s) of lagged consumption plugged into the current price regime, *i.e.*,

$$z_{i,t} = p_{i,t}\left(q_{t-k}\right)$$

Pro: Matches multi-tiered price structure and removes mechanical price-quantity link.
Con: Can introduce more noise.

In practice: We use quantity lags 10–14 to “vote” whether the household is on the first or second tier, *i.e.*,

$$v_{i,t} = \frac{1}{5} \sum_{k=10}^{14} \mathbb{1}\left\{q_{i,t-k} > \bar{A}_{i,t}\right\}$$

where $\bar{A}_{it}$ is household *i*’s baseline allowance in time *t*.

The instrument:

$$z_{i,t} = \mathbb{1}\left\{v_{i,t} \leq 0.5\right\} \times p_{i,t}^{\text{base}} + \mathbb{1}\left\{v_{i,t} > 0.5\right\} \times p_{i,t}^{\text{excess}}$$