Title
Demand for Three Environmentally Friendly Goods: Plant Based Meat, Clean Air, and Efficient Light Bulbs

Author
Gordon, Hal Giuliani

Publication Date
2022

Peer reviewed|Thesis/dissertation

Demand for Three Environmentally Friendly Goods: Plant Based Meat, Clean Air, and Efficient Light Bulbs

by

Hal Gordon

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Agriculture and Resource Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Meredith Fowlie, Chair
Professor Sofia Villas-Boas
Professor Miguel Villas-Boas

Fall 2022

Demand for Three Environmentally Friendly Goods: Plant Based Meat, Clean Air, and Efficient Light Bulbs

Abstract

Demand for Three Environmentally Friendly Goods: Plant Based Meat, Clean Air, and Efficient Light Bulbs

by

Hal Gordon

Doctor of Philosophy in Agriculture and Resource Economics

University of California, Berkeley

Professor Meredith Fowlie, Chair

Chapter 1

Meat and agriculture is a major source of greenhouse gas releases, and beef is particularly emissions heavy. Plant-based meat (PBM) is billed as a new food that could overturn the beef market and eventually transition many meat eaters to a far more environmentally friendly option. Currently, we have very little information about how much meat is being substituted by PBM. This paper uses proprietary data from a nationwide grocery chain to create a very large sample of households who have bought PBM at least once. With this dataset, I am able to draw fine conclusions about what attributes are related to purchasing and repurchasing PBM. I find that buying and more importantly, rebuying PBM is associated with having previously bought less meat and more meat substitutes. In addition, the people entering the PBM market are no more likely to have bought meat than those who first started buying it, suggesting PBM is struggling to expand its reach to those who could most easily switch away from real meat. In addition, because of how promotional pricing is determined at this nationwide chain, I am able to run event study regressions to test the theory that PBM has is a robust substitute for beef in grocery stores. In these regressions, I find little evidence for switching between meat and PBM.

Chapter 2

(Co-authored with Scott Kaplan) This project leverages a unique setting to study the effects of air pollution on a market good people consume for recreational purposes: tickets to National Football League (NFL) games posted on a popular, online secondary marketplace. Our initial findings suggest that an increase in the AQI does not lead to a statistically significant change in listed ticket prices (in fact, we find a slightly positive estimate that is likely the effect of unaccounted for noise). We also determine that there was no statistically significant impact on the number of tickets listed on the marketplace for these games.

Chapter 3

The so-called energy efficiency gap that describes how households apparently undervalue investments in energy efficient appliances and capital improvements has long vexed policy makers. In 2007, the US Congress attempted to help close this gap by establishing a 30% efficiency mandate on all existing incandescent light bulbs. The aim of the law was to eliminate cheap, inefficient bulbs in favor of more expensive, highly efficient compact florescent lamps (CFLs) or yet to be produced more efficient halogen incandescent bulbs, but the ultimate effect of the mandate was unclear at the time it was passed. Using heterogeneity in the roll-out of the mandate between California and the rest of the country, I use a panel of light bulb sales data from a nationwide discount store retailer to examine the effects on the market the mandate had. I find that the mandate increased the price of incandescents (along with their efficiency) but had little effect on the demand for CFLs. When certain incandescents were made unavailable I find strong evidence that customers switched to different types of incandescents or, in some instances, reduced overall light bulb purchases. The mandates clearly incrementally increased the efficiency of light bulbs purchased, but did not switch the nation en-mass to high efficiency CFLs.

Contents

List of Figures

List of Tables

Acknowledgments

I want to thank: my advisors, Meredith Fowlie, Sofia Villas-Boas, and James Sallee. They were there for me and very understanding when COVID put a wrench in all of our plans. My wife, Kelsey, who supported me in all kinds of ways for all these years, especially when I wanted to quit. My parents and sister who have been great cheerleaders. My entire cohort, who became a great family and lifelong friends. My friend Jim, who made me feel like San Francisco was my home. All of my friends at the Coliseum, who after a day of frustrating coding built me back up with their kind words and propensity to repeat back whatever silly cheer I yelled at them.

Chapter 1

Predicting the Demographics of Plant Based Meat Customers

1.1 Introduction

Plant-based meat (PBM) is a brand new food category that has attracted widespread media and investor attention since its introduction in 2017. In contrast to other meat substitutes (tofu, veggie burgers), plant-based meat (PBM) products are being marketed as indistinguishable in taste and appearance to meat. Producers of plant-based meats aim to compete more directly with, and eventually replace, meat.

The appearance of these items (still mostly confined to ground beef and sausages) beginning in 2017 has drawn a significant amount of interest from environmentalists, conservationists, and animal welfare advocates, as they hope it will result in lower demand for beef and wilderness conversions to pasture required to raise it . The potential is massive. Meat and dairy account for 15% of global greenhouse gas emissions (Gerber et al. 2013), and the emissions produced by a kilogram of beef can be as much as ten times higher than pork or chicken and twice as high as the second worst offender, lamb (Poore and Nemecek 2018). For those concerned about biodiversity, conversion of land to grazing pasture or feed agriculture is the leading cause tropical forest deforestation in much of the world (Henders, Persson, and Kastner 2015).

There is a lack of political will to support the imposition of new taxes on meat products. As such, the only viable strategy (beyond appealing to emissions reductions or animal welfare) for plant-based meat to gain market share from traditional animal-derived meat is to compete on price and taste. Some futurists have suggested that technological advancements may make this inevitable in the future (Huang, Gordon, and Zilberman 2020), but currently PBM is far more expensive. While alternative protein sources have gained market share in recent years, it remains to be seen whether they will be able to replace their animal-based counterparts in a significant portion of American's diets.

Researchers have sought to better understand the potential for PBM, but because it is so new, much work has been limited to stated preference surveys, which rely on hypothetical choices. Researchers have generally found that even if offered large price reductions, the market share for PBM is likely to stay below 20% of ground meat (Van Loo, Caputo, and Lusk 2020, Carlsson, Kataria, and Lampi 2022). Research that takes advantage of retail scanner data has sought to examine substitution patterns, but without individual level data, can only say limited things about who the customers are (Zhao et al. 2022).

Many of the existing stated preference studies identify those already predisposed to meat alternatives as the most likely to purchase PBM, and also find a not insignificant amount of the demand comes from those being pulled into the meat market (Van Loo, Caputo, and Lusk 2020), yet researchers have also found that most PMB purchasers (80 to 90%) also bought meat or ground meat.

Understanding both who PBM purchasers are in the real world and how their consumption is affected is critical to understanding if PBM has a real potential to replace beef and other meats, or if PBM is simply a new category of meat alternative that is mainly competing with other kinds of low emissions foods like veggie burgers, tofu, or chicken. As data availability has increased some researchers have been able to use household scanner data to begin to answer these questions, but there are limited sample sizes due to the small size of PBM market share. In particular, this work has shown us that PBM purchasers did not reduce their ground beef purchases after they first bought PBM, and that just more than half of PBM purchasers bought it again (Neuhofer and Lusk 2022).

This paper has the advantage of proprietary data from a nationwide grocery chain that goes beyond home scanner data. With this data, I am able to draw a very large sample of households who have bought PBM at least once and have a long history of their purchases before and after. With this dataset, I am able to draw more fine conclusions about what attributes are related to purchasing and repurchasing PBM, as well as if PBM is able to spread to customers who are more likely to be substituting away from meat. In addition, because of how promotional pricing is determined at this nationwide chain, I am able to run event study regressions to test the theory that PBM has is a robust substitute for beef in grocery stores.

1.2 Data

The data for this study are proprietary data from a single nationwide grocery store chain, which operates in a number of different regions and under a number of different store or dividion names. Over 90% of sales across all brands are linked to an individual membership card that is theoretically assigned to a single household, although some people may share their membership card. When a customer applies for a membership card, some but not all have provided certain information (phone number and address) that has allowed the grocery chain to match them with data from a credit company who provides estimates of the head of household's age and race[1], and the size and income of the household.

Prior studies using person level scanner data have been limited by the number of households that have ultimately purchased plant based meat. In order to ensure a large sample size, every household that purchased plant-based meat from its first limited introduction in 2017 to the end of 2019 was included, and a random sample of all households who did not purchase PBM in the same period was included as a comparison.

While PBM was only first introduced in limited stores in mid 2017, all purchases since the beginning of 2016 are included in the data. This allows me to characterize households by their purchases in prior periods, as well as flexibly examine what types of purchases could ultimately predict who will buy PBM and who will rebuy it, with the ultimate aim of trying to better understand if PBM are attracting the types of customers who are likely to substitute it for real meat.

About 40% households did not have a full credit data match, either because they did not share enough data when they signed up for their membership card to be linked in the credit company's database, or because one of the demographics estimates was missing for other unknown reasons. Households whose first trip was later than the 12th week of 2016 and whose last trip was before the 40th week of 2019 were also dropped. The spread of average trips per week per household between those that had and those who hadn't bought PBM were found to be fairly different. This makes intuitive sense, as households that recorded few trips were more unlikely to have purchased any specific product item coded as a UPC. As a result the sample was trimmed to make the distribution of trips more similar: only households with more than 0.85 trips per week and less than 5.85 trips per week were included. After all these cuts, we were left with 134,167 households who had bought PBM at least once and 60,269 households that had never bought PBM.

In order to test if political partisanship has had an effect on demand for PBM, an additional demographic variable was created by connecting the zipcode of the most frequented store by each household to precinct level voting data from the 2016 presidential election compiled by the New York Times(Upshot 2018). Because the precinct that a store is in might be different from the surrounding area, the percentile Clinton vote of each precinct's surrounding area is used, which the New York Times describes as "a measure based on the

[1]Race estimates are expressed as probabilities, so each household has a value between 0 and 1 for each race.

choices of the nearest 100,000 voters, as well as those within a 10-mile radius".

Plant-based meat beef (as opposed to plant-based meat sausage, which were not introduced until mid 2019) were first rolled out in select areas in mid 2017, so the entire 2016 year is used as a pre-period to establish what households spent their money on before PBM existed. Five different pre-period variables were created due to assumptions of importance. Four variables were defined as the percent of expenditures in the pre-period that were spent on (1) meat, (2) ground beef, (3) veggie burgers, and (4) tofu. In addition, a household specific low meat indicator was defined as equal to one if a household spent less than 5% of expenditures in the pre-period on any meat.

In order to expand the potential important relationships of pre-period purchases to not just the four goods of interest, but to any goods, two levels of UPC categorization used by the grocery chain were used. All UPCs are organized into 57 'Groups' and the more specific 459 'Categories' and those divisions were used for percent of expenditures in the pre-period calculations.

1.3 Variables of Interest

There are three outcomes of interest: buying plant-based meat, rebuying PBM, and the changing demographics of who is buying PBM for the first time. Simply buying PBM is of obvious interest, but rebuying it may be even more important. If PBM is going to replace real meat, it will have to become a regular part of customers' market basket. Simply knowing who the buyers and rebuyers of PBM are will not in itself provide a causal estimate of how much meat we expect PBM to replace, but it can tell us how optimistic we are about its ability to outcompete meat. Even the most optimistic PBM evangelists do not expect PBM to make a large dent in meat sales this early in its existence, but they do often cite the fact that 80 to 90% of PBM customers also eat meat. The implication is that PBM is not just for vegetarians, but will slowly muscle out meat in these customer's diets. While this simple fact has been shown to be true, a much closer look at who is buying and rebuying PBM will tell us more about its potential as a meat killer.

Because I observe grocery sales up until the last day of 2019, however, some customers have far longer to rebuy from their first purchase than households that bought near the end of 2019. If I used a simple 1/0 variable of rebuy, then I would find that early first time buyers were far more likely to have rebought, but as will be discussed in the next paragraph, there is reason to believe that later first time buyers may be different from earlier ones, so if this timing issue is not corrected for, we may mistake these changes in who is in the market for PBM as related instead to who is rebuying (or not rebuying) it. Therefore, I limit the rebuy decision to 12 weeks after the first buy for each household, and I drop all households who bought their first PBM item within 12 weeks of the end of 2019.

Less obvious is why I care about the timing of when households first bought PBM. The proponents of PBM market it as not a better veggie burger, but as a whole new category of food, so the data are witnessing how a brand new option enters the grocery market. In mid 2017, when PBM was first offered for sale on grocery store shelves, it would not be surprising if many of the first time customers were highly knowledgeable about meat and animal alternatives and that those purchases were unlikely to be replacing much meat. That, however, is much less important than how the market evolves. If more and more meat eaters are entering the market as it matures, then PBM is reaching those who are more likely to replace meat in their diet. Seeing evidence of PBM going mainstream the meat eaters would be encouraging for future growth.

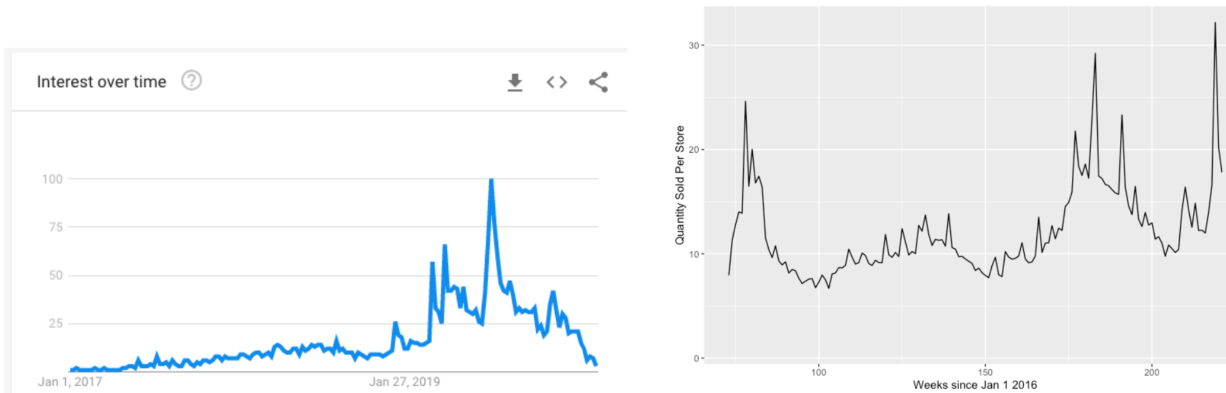
In order to test how the market of first time buyers is changing in a comparable way to buy and rebuy, I created a 1/0 variable of early(1) and late(0). The cut off is set at April 29th, 2019 (the first day of the 18th week of 2019, or the 174th week in the sample) and while this may seem partly arbitrary, it closely matches when the public became much more aware of PBM. This was close to when Burger King first began running national ads for its Impossible Whopper, which attracted an unprecidented amount of attention for PBM. In addition, by this time, PBM had been available in all stores across this chain of grocery stores for at least 6 months, but this week marked the bgining of a nationwide increase in demand for PBM in this grocery chain.

We can see both of these trends in figure 1.1. On the left is a Google Trends search of Plant Based Meat for 2017 to 2019, showing the large spikes in interest in the spring of 2019. On the left, I show the quantity of Plant Based Meat sold per store (with positive sales) per week. At the beginning, sales per store are high, but that may reflect the limited geographies that the item was made available in. As more and more stores started selling PBM, sales per store leveled out at a much lower level, until the spring of 2019, when sales began to rise again as nationwide interest in PBM picked up. In the next section, I will use week of first purchase (rather than a binary early/late variable) as an alternate that can pick up more variation.

Table 1.1 displays all the covariables discussed and how they relate to the three variables of interest, as well as t stats and p values. With such large sample sizes, the covariates on the variables of interest are unsurprisingly almost always statistically different between the two groups of households.

Of particular interest for the buy / no-buy groups, those who bought PBM were more likely to be younger, to have a higher income, and to be shopping in a more Clinton heavy area. The extra variables created from the pre-period market basket are also interesting, if not completely surprising. While households were more likely to be "low meat" households

Figure 1.1: Google Trends for Plant Based Meat and Demand per Store



that spent less than 5% of their pre-period budget on meat, the difference was only a half percentage point. This supports the often repeated fact that PBM customers also buy meat. However, they still bought less meat, with a larger gap in their average market basket raw percent devoted to meat in the pre period. In addition, buyers of PBM bought about 3 times as much veggie burgers and tofu. So while PBM buyers did, on average, also buy meat, they bought less of it, and also bought a lot more traditional meat replacements. Because the number of buyers of PBM is oversampled, I cannot say what percent of all customers have bought it at least once.

The early/late comparison and rebuyers in the second and third section of table 1.1 only include those who bought PBM once, so the difference in these groups would not necessarily be the same as between the buyers and no buyers. The demographics of rebuyers were only slightly different: skewing slightly younger and shopping in slightly more Clinton voting areas. However, rebuyers were more likely to be "low meat" purchasers and had spent more on traditional replacements and a little less on meat in the pre-period. That seems to infer that PBM was more likely to catch on with customers who were already interested in replacing meat in their diets. Finally, it's important to note that less than a third of customers rebought PBM in the three months after first trying it. Because I have trimmed the sample to only include households that make regular trips (a minimum of 0.85 trips a week), each customer had plenty of opportunities to buy PBM again, and this shows that PBM has not yet converted many of the people who try it to become regular purchasers.

There are also some clear patterns in how earlier first time buyers compared to later ones. Early buyers were younger and less wealthy (although still wealthier than those who never bought at all). Earlier buyers were more likely to be 'low meat', bought less meat and ground beef, and bought more meat replacements. This suggest that later purchasers were less predisposed to buying tofu or veggie burgers, so were better candidates to be replacing thier meat intake, but comparing late buyers to those who never bought PBM is more mixed. Later buyers are slightly less likely to have bought low amounts of meat than non PBM customers, but on average, they still spent less total on meat and ground beef and spent more on tofu and veggie burgers. This means that the most meat heavy consumers are still holding out on trying PBM.

1.4 Lasso Models

While somewhat informative, the covariates in table 1.1 do not include the vast majority of the information known about the households, namely their purchasing history. With such a large number of observations and covariates (there are 459 categories of food), it's a good assumption that hundreds of variables will have significant differences between buyers and non-buyers. That does not mean all these variables are important. In order to see how well we can predict the variables of interest, we can load all of those variables into a linear regression and examine the R^2. This will likely result in a large amount of overfitting, and with hundreds of variables being found to have significant relationships with the outcomes, little practical value.

For this reason, I will implement those regressions using the least absolute shrinkage and selection operator (Lasso), a common machine learning algorithm that automatically selects a model so that many coefficients go to zero. It differs from OLS by minimizing the sum of squared residuals plus the absolute value of the coefficients times a penalty. Lassos test each variable in a regression for whether it has additional explanatory power above the other variables (Tibshirani 1996). This is particularly useful in this case, since we are not trying to find causal inferences, only particularly robust relationships between the demographics and purchase history of these customers, and their decisions about PBM. It is important to try Lassos using a few different specifications, since they can be highly sensitive to specification(Mullainathan and Spiess 2017). For each variable of interest, I run an OLS regression and a Lasso regression (using either the 459 categories or 57 groups) of the following form:
$V_n = \beta C_n + \alpha D_n + \delta A_n + \epsilon_n$

Where V_n is the variable of interest, C_n is the vector of proportion of spending coming from either all groups or categories in the pre-period, D_n are the demographic variables, and A_n are the five extra pre-period variables that I hypothesised would be important before running any of these regressions.[2] Because Lasso regressions can be sensitive to specifications, I ran each regression and Lasso using just the pre-period variables, using the pre-period vari-

[2]Because meat is itself a single defined food group, I drop it from my extra variables in models using groups

Table 1.1: Household Means

	Bought / Didn't Buy PBM				Early / Late Adopters				Rebuy within 3 Months			
	Buy	No Buy	T Stat	P Value	Early	Late	T Stat	P Value	Rebuy	No Rebuy	T Stat	P Value
Demographic Vars												
N HHs	134, 167	60, 269			62, 505	71, 662			42, 428	91, 739		
Trips	382.15	309.63	81.545	0.000	386.29	378.55	7.166	0.000	394.83	376.29	15.897	0.000
Age	53.75	55.11	−16.614	0.000	53.20	54.23	−11.824	0.000	53.49	53.88	−4.210	0.000
Income	110.80	96.87	41.950	0.000	106.56	114.50	−20.643	0.000	110.99	110.72	0.657	0.511
Prob White	0.68	0.70	−12.485	0.000	0.68	0.69	−4.962	0.000	0.68	0.69	−4.393	0.000
Prob Black	0.06	0.06	4.050	0.000	0.05	0.07	−27.268	0.000	0.07	0.06	6.743	0.000
Prob Asian	0.13	0.11	28.263	0.000	0.13	0.12	8.620	0.000	0.13	0.13	−0.936	0.349
HH Size	3.55	3.35	20.363	0.000	3.52	3.58	−5.412	0.000	3.52	3.56	−3.633	0.000
Area 18 Clinton	73.45	66.21	65.887	0.000	73.12	73.74	−5.543	0.000	73.85	73.26	4.905	0.000
Region: Denver	0.04	0.07	−32.272	0.000	0.04	0.04	−0.456	0.648	0.03	0.04	−8.946	0.000
Region: Eastern	0.11	0.10	6.098	0.000	0.04	0.17	−86.184	0.000	0.12	0.10	9.670	0.000
Region: Intermountain	0.00	0.01	−14.311	0.000	0.00	0.00	−3.365	0.001	0.00	0.00	−4.184	0.000
Region: Northern CA	0.31	0.23	38.617	0.000	0.36	0.27	34.870	0.000	0.32	0.31	1.380	0.168
Region: Portland	0.04	0.07	−26.762	0.000	0.04	0.04	3.757	0.000	0.03	0.04	−2.179	0.029
Region: Seattle	0.04	0.07	−25.853	0.000	0.04	0.04	13.478	0.000	0.03	0.04	−9.316	0.000
Region: Southern	0.04	0.07	−48.538	0.000	0.04	0.04	0.262	0.793	0.03	0.04	−9.206	0.000
Region: Southern CA	0.04	0.07	54.540	0.000	0.04	0.04	12.816	0.000	0.03	0.04	5.087	0.000
Region: Southwest	0.04	0.07	−18.388	0.000	0.04	0.04	−14.005	0.000	0.03	0.04	3.216	0.001
Extra Vars												
Perc Low Meat	18.94%	18.23%	3.715	0.000	20.40%	17.67%	12.680	0.000	22.72%	17.19%	23.195	0.000
Basket %-Meat	11.83%	13.02%	−27.778	0.000	11.64%	12.00%	−8.507	0.000	11.07%	12.18%	−24.504	0.000
Basket %-Grd Bf	1.31%	1.66%	−32.689	0.000	1.29%	1.32%	−2.945	0.003	1.20%	1.36%	−16.167	0.000
Basket %-V Burg	0.10%	0.03%	44.703	0.000	0.12%	0.08%	13.756	0.000	0.14%	0.08%	19.399	0.000
Basket %-Tofu	0.08%	0.03%	38.864	0.000	0.10%	0.07%	14.768	0.000	0.11%	0.07%	15.813	0.000

ables and demographics, using the pre-period variables and the extra pre-period variables, and finally one with all three sets of variables. The R^2 and the number of nonzero coefficients are reported in the tables below. I also split the sample in half, and reran the regressions and Lassos on half the sample (the training data) and then used those coefficients to make predictions on the testing data and then calculated the root mean squared error (RMSE). A higher RMSE is indicative of more overfitting, something we should be worried about with this amount of coefficients.

Lasso frees the researcher to try many different types of regression specifications. I tried Lassos that used quadratics on the variables, and regressions that interacted the demographics with all of the pre-period variables. I also ran some regressions using pre-period market basket proportions at the UPC level, which expands those variables from 459 to hundreds of thousands. Because of the very large sample size (nearly two hundred thousand households who average about 350 trips), these Lasso regressions quickly became computationally challenging, so they were used to inform how I could improve the smaller Lassos reported in tables 1.2 and 1.3. The main takeaways from these unreported Lassos were that interactions between the regions were very important. This was not surprising, since PBM was rolled out at different times in different regions, and the price was set at the region level. The first two regions that had access to PBM were the Northern and Southern California regions. In order to examine these interactions in a less computationally difficult way, I reran all the models in tables 1.2 and 1.3 on just the households that shopped in those two regions (tables 1.4 and 1.5). In order to improve the predictive power fo the models, I use an additional variable of interest, week of first purchase, in place of the the 1/0 early/late variable(tables 1.6 and 1.7).

Lasso Results

The OLS and Lasso results are at first blush disappointing. The highest R^2s are in table 1.2 for the buy/don't buy decision. There, the OLS models have an R^2 of 0.16, which is not particularly high. The RMSE are very close to 0.5 (the approximate maximum value given a 1/0 outcome), which (unsurprisingly) suggests loading in almost 500 variables into an OLS creates a lot of overfitting. The Lassos have slightly lower R^2s but the RMSE is more respectable, meaning some (but not all) of the overfitting is dealt with. Unfortunately, the Lassos only collapse the coefficients to zero of about have the category variables. This means we are not able to single out a handful of traits that are much more important in predicting the buy/don't buy decision.

The models on early and late adoption and rebuying within 3 months have even lower R^2s, but are at least able to reduce the number of non-zero coefficients a bit better. The two are no doubt related, with a lower R^2 meaning that the models predict fewer and fewer decisions, and fewer decisions successfully predicted means the Lasso can reduce more of the coefficients to zero. What is true for all three variables of interest is that the 5 extra

Table 1.2: OLS and Lasso Models using Pre Sale Data by Category

	Bought / Didn't Buy PBM				Early / Late Adopters				Rebuy within 3 Months			
	Cat	Cat/ Demos	Cat/ Extras	Cat/ Demos/ Extras	Cat	Cat/ Demos	Cat/ Extras	Cat/ Demos/ Extras	Cat	Cat/ Demos	Cat/ Extras	Cat/ Demos/ Extras
OLS Models												
r^2	0.161	0.161	0.161	0.161	0.077	0.077	0.077	0.077	0.031	0.031	0.031	0.031
RMSE of test data	0.497	0.498	0.497	0.499	0.518	0.519	0.518	0.519	0.482	0.482	0.482	0.482
Lasso Models												
r^2	0.148	0.155	0.151	0.157	0.068	0.071	0.068	0.071	0.022	0.022	0.023	0.022
RMSE of test data	0.428	0.427	0.428	0.426	0.482	0.482	0.482	0.481	0.469	0.469	0.469	0.469
Non-Zero Coeffs												
Cats (out of 459)	210	215	232	230	92	106	105	106	50	41	59	42
Demos (out of 15)		14		14		13		13		6		6
Extras (out of 5)			4	4			0	1			5	5

Table 1.3: OLS and Lasso Models using Pre Sale Data by Group

	Bought / Didn't Buy PBM				Early / Late Adopters				Rebuy within 3 Months			
	Group	Group/ Demos	Group/ Extras	Group/ Demos/ Extras	Group	Group/ Demos	Group/ Extras	Group/ Demos/ Extras	Group	Group/ Demos	Group/ Extras	Group/ Demos/ Extras
OLS Models												
r^2	0.122	0.122	0.122	0.122	0.065	0.065	0.065	0.065	0.018	0.018	0.018	0.018
RMSE of test data	0.486	0.489	0.488	0.49	0.513	0.514	0.513	0.515	0.468	0.468	0.469	0.469
Lasso Models												
r^2	0.104	0.115	0.11	0.121	0.057	0.063	0.059	0.064	0.012	0.013	0.014	0.016
RMSE of test data	0.439	0.436	0.437	0.435	0.484	0.483	0.484	0.483	0.463	0.462	0.462	0.462
Non-Zero Coeffs												
Groups (out of 57)	44	48	46	50	32	39	29	33	19	23	15	25
Demos (out of 15)		15		15		13		13		8		9
Extras (out of 4)			4	4			3	3			4	4

Table 1.4: OLS and Lasso Models using Pre Sale Data by Category, CA only

	Bought / Didn't Buy PBM				Early / Late Adopters				Rebuy within 3 Months			
	Cat	Cat/ Demos	Cat/ Extras	Cat/ Demos/ Extras	Cat	Cat/ Demos	Cat/ Extras	Cat/ Demos/ Extras	Cat	Cat/ Demos	Cat/ Extras	Cat/ Demos/ Extras
OLS Models												
r^2	0.099	0.099	0.099	0.099	0.028	0.028	0.028	0.028	0.03	0.03	0.03	0.03
RMSE of test data	0.444	0.445	0.444	0.446	0.507	0.507	0.507	0.507	0.481	0.481	0.481	0.481
Lasso Models												
r^2	0.081	0.087	0.083	0.088	0.014	0.02	0.017	0.021	0.019	0.018	0.017	0.019
RMSE of test data	0.409	0.408	0.409	0.407	0.496	0.495	0.496	0.495	0.47	0.47	0.47	0.47
Non-Zero Coeffs												
Cats (out of 459)	139	137	150	135	47	62	72	77	28	24	13	31
Demos (out of 8)		7		7		6		6		1		1
Extras (out of 5)			4	4			0	0			4	4

Table 1.5: OLS and Lasso Models using Pre Sale Data by Group, CA only

	Bought / Didn't Buy PBM				Early / Late Adopters				Rebuy within 3 Months			
	Group	Group/ Demos	Group/ Extras	Group/ Demos/ Extras	Group	Group/ Demos	Group/ Extras	Group/ Demos/ Extras	Group	Group/ Demos	Group/ Extras	Group/ Demos/ Extras
OLS Models												
r^2	0.063	0.063	0.063	0.063	0.015	0.015	0.015	0.015	0.015	0.015	0.015	0.015
RMSE of test data	0.435	0.437	0.436	0.438	0.501	0.503	0.502	0.503	0.47	0.47	0.471	0.471
Lasso Models												
r^2	0.044	0.054	0.049	0.058	0.006	0.012	0.007	0.013	0.008	0.008	0.011	0.011
RMSE of test data	0.416	0.414	0.415	0.413	0.499	0.497	0.498	0.497	0.466	0.465	0.465	0.465
Non-Zero Coeffs												
Groups (out of 57)	39	41	38	43	35	33	31	27	12	12	12	13
Demos (out of 8)		7		7		6		6		0		0
Extras (out of 4)			4	4			3	3			4	4

Table 1.6: OLS and Lasso Models Estimating Week of First Buy

	Categories				Groups			
	Cat	Cat/ Demos	Cat/ Extras	Cat/ Demos/ Extras	Group	Group/ Demos	Group/ Extras	Group/ Demos/ Extras
OLS Models								
r^2	0.095	0.095	0.095	0.095	0.081	0.081	0.081	0.081
RMSE of test data	38.341	38.579	38.343	38.581	37.907	38.213	37.951	38.252
Lasso Models								
r^2	0.074	0.089	0.076	0.088	0.06	0.075	0.063	0.078
RMSE of test data	35.419	35.172	35.391	35.159	35.673	35.348	35.621	35.348
Non-Zero Coeffs								
Cats (out of 459 / 57)	69	90	103	75	18	22	17	20
Demos (out of 8)		13		13		13		13
Extras (out of 5 / 4)			0	1			3	3

Table 1.7: OLS and Lasso Models Estimating Week of First Buy, CA Only

	Categories				Groups			
	Cat	Cat/ Demos	Cat/ Extras	Cat/ Demos/ Extras	Group	Group/ Demos	Group/ Extras	Group/ Demos/ Extras
OLS Models								
r^2	0.051	0.051	0.051	0.051	0.034	0.034	0.034	0.034
RMSE of test data	39.076	39.352	39.081	39.356	38.505	38.89	38.536	38.915
Lasso Models								
r^2	0.022	0.041	0.024	0.036	0.003	0.028	0.009	0.029
RMSE of test data	37.952	37.613	38.016	37.599	38.271	37.859	38.248	37.825
Non-Zero Coeffs								
Cats (out of 459 / 57)	51	37	76	10	5	14	15	10
Demos (out of 8)		7		4		7		7
Extras (out of 5 / 4)			0	0			3	3

variables that I created by themselves offer very little collective improvement to the explanatory power of the models. Indeed, unreported OLS models that exclude the 459 category variables but include just the 5 extra variables have very low R^2s across the board. The demographic variables (partly due to, but not entirely due to, the importance of region) are far more important.

In comparison to the category models, the group models have yet lower R^2s (not too surprising given the far lower number of variables) and at first blush, manageable numbers of non-zero coefficients. However, as a percentage of covariates, those with non-zero coefficients are actually higher.

In an attempt to further reduce the large number of non-zero coefficients, and using the knowledge from limited runs of Lassos that included large numbers of interactions, I reduced the sample to just those households in California (the two regions that the grocery chain first made PBM available). As the sample size is reduced further, the R^2s continue to go down, but at least for the rebuy decision, the Lasso is able to collapse the non-zero variables down to a reasonable number(tables 1.4 and 1.5).

In order to better examine the timing of when households first tried PBM, I ran the same models, but predicted week of first purchase as an integer. Tables 1.6 and 1.7 show higher R^2s, although with pretty significant dropoffs for the Lasso models, especially when just restricted to California households. Still, these models are able to pare down the number of non-zero coefficients while also improving the explanatory power, so they are preferred.

With the knowledge that these models fail to explain much of the variation, it is still useful to look at what covariates were non-zero in the Lassos for hints about what is important. To that end, I have listed the non-zero coefficients for the 3 month rebuy California only Lassos in table 1.8 and for the week of first buy California only in table 1.9. These tables only list non-zero category coefficients that were non-zero in at least 3 of the 4 specifications. Some categories require a bit of explanation. 'Mainstream White', a negative rebuy category, encompasses most milk. ISB rolls, snacks, and desserts, all negatively associated with rebuying, are items from the grocery stores' own bakery. 'Asian Healthy', positively related to rebuys and negatively related to weeks, includes tofu, and also includes noodles, kimchi, and dumplings. 'Refrigerated Premium Drinks' are bottled kombuchas, coffees, and expensive juices. 'Meal Compnts Frozen Prepared Food' are frozen side dishes, mostly pastas. 'Refrigerated (Retail Packs) is non-frozen, non-canned seafood in retail packaging (smoked salmon, fresh shellfish, etc.). 'Meat/Dairy Alternative' are veggie burgers, tofurkey, soyrizo, and vegan cheese.[3]

[3] I do not understand what "Container Deposit Produce" is, but the amount of expenditures on it is very small

Table 1.8: Non-zero Coefficients: Rebuy 3 mths, CA only

Negatively Related to Rebuys			Positively Related to Rebuys		
Categories (% basket in pre-period)	# of Non-Zeros	Sum of Coeffs	Categories (% basket in pre-period)	# of Non-Zeros	Sum of Coeffs
Canned/Jar Pineapple	3	-1.607	Meat/Dairy Alternative	4	9.058
Mainstream White	4	-1.594	Meal Compnts Frozen Preped Food	4	5.532
Cream	4	-1.344	Finfish All Other - Wild (Bulk)	3	1.941
Bacon	4	-1.211	Asian Healthy	3	1.874
Lettuce	3	-1.129	Organic Vegetables	4	1.625
Carbonated Soft Drinks	4	-0.995	Avocados	4	1.173
Beef, Choice, No Brand	4	-0.867	Rice Side Dishes	4	1.031
Pork, Natural	3	-0.785	Shelf Stable Pasta & Pizza Sauce	3	0.627
Isb Rolls	3	-0.542	Deli Dips And Salsa	4	0.567
Service Meat	3	-0.330	Refrigerated Premium Drinks	4	0.451
Isb Snacks	3	-0.226	Refrigerated (Retail Packs)	3	0.341
Isb Desserts	4	-0.124			
Demos & Extras	# of Non-Zeros	Sum of Coeffs	Demos & Extras	# of Non-Zeros	Sum of Coeffs
Meat	2	-0.268	Tofu	2	3.025
Ground Beef	2	-0.238	Low Meat	2	0.018
			Region: So Cal	2	0.004

Taken holistically, table 1.8 does not give a lot of information on who is most likely to rebuy PBM. Milk, bacon, beef, meat, pork, and ground beef are all predictive of not rebuying PBM, while meat and dairy alternatives, seafood, tofu, kombucha and expensive produce are related to rebuying. This table tells a clear story that customers who are more likely to incorporate plant based meat into their diets are much less likely to have had a diet high in the ground beef that PBM cheerleaders hope to be replacing.

But maybe that is changing? Table1.9 shows the non-zero coefficients related to the first week of purchase. Again, there is a clear story to be told about the early adopters of PBM. They are more likely to have bought meat and dairy alternatives, tofu, and organic vegetables. This fits the story that early adopters were more interested in and better informed about PBM. However, the positive variables don't tell a very easy to imagine story. Various types of candy, fruits, yogurt, and deli cheese seem like a pretty normal market basket. What is missing is meat. If the industry was hoping that over time more and more 'meat eaters' were trying PBM, these Lassos don't support that story. Getting Impossible burgers into Burger King did not spur heavy meat purchasers to finally try PBM in grocery stores.

Table 1.9: Non-zero Coefficients: Week of First Buy, CA only

Negatively Related to Week of 1st Buy			Positively Related to Week of 1st Buy		
Categories (% basket in pre-period)	# of Non-Zeros	Sum of Coeffs	Categories (% basket in pre-period)	# of Non-Zeros	Sum of Coeffs
Container Deposits Produce	4	-3684.839	Kosher Specialty Foods	3	237.640
Meat/Dairy Alternative	4	-1042.630	Pears	3	150.309
Asian Healthy	4	-578.837	Apples	4	137.287
Meal Compnts Frozen Preped Food	4	-412.981	Deli Cheese	4	137.139
Single Serve Sweet Baked Goods	3	-399.177	Condensed Multi-Serve Soup	3	122.057
Frozen Potatoes & Onions	4	-256.368	Xmas Candy, Gum & Mints	3	121.324
Organic Frozen Vegtables	3	-154.388	At Home Crackers	4	117.226
Organic Vegetables	4	-107.544	Laundry Pre-Treatment	3	81.650
Canned Chili	3	-106.978	Beverage Ice	3	79.031
Specialty Oils - Flavored & Unflav	3	-84.275	Blueberries	3	69.211
Dry Beans	3	-65.393	Tomatoes	3	58.944
Frozen Snack Prepared Foods	3	-62.191	Refrigerated Yogurts	4	50.905
Pizza Frozen Prepared Foods	3	-23.601	Baking Additives	3	47.871
			Mandarins	3	39.199
			Candy/Gum/Mints (In-Line)	3	35.041
			Cold Cereal	3	23.862
Demos & Extras	# of Non-Zeros	Sum of Coeffs	Demos & Extras	# of Non-Zeros	Sum of Coeffs
Probability Black	2	-0.008	Region: So Cal	2	18.750
			Age	2	0.023
			HH Income	2	0.020
			Clinton vote 2016	1	0.009
			Probability Asian	1	0.004
			Probability White	1	0.003

Lasso Summary

It seems that the information in these grocery data, while exceptionally rich and large, still struggles to predict the majority of PBM decisions. This is not surprising. I ran a similar buy/no buy model on tofu, which has a similar market share and is a similar product, and got very similar R^2s and non-zero variable results 1.10. Because of this, I reject the idea that PBM is particularly hard to predict, but just that it is hard to accurately predict any of these variables of interest in a grocery context. Data from a much more limited choice set may be preferable, like from a Burger King. However, the results that are interpretable do not have much good news for PBM. Even among those who try it at least once, those who are more likely to eat meat are less likely to rebuy it, and PBM has not been more successful at being attractive to meat eaters who have held out, even after its profile was raised greatly and it was sold at fast food restaurants. From these data, I conclude that PBM has a lot longer to go on making itself more attractive to the kinds of customers who are most likely to be substituting away from beef.

Table 1.10: OLS and Lasso Models Predicting Buying Tofu

	Categories				Groups			
	Cat	Cat/ Demos	Cat/ Extras	Cat/ Demos/ Extras	Group	Group/ Demos	Group/ Extras	Group/ Demos/ Extras
OLS Models								
r^2	0.165	0.165	0.165	0.165	0.133	0.133	0.133	0.133
RMSE of test data	0.518	0.519	0.519	0.52	0.503	0.505	0.511	0.513
Lasso Models								
r^2	0.152	0.157	0.157	0.161	0.09	0.099	0.122	0.131
RMSE of test data	0.445	0.444	0.444	0.443	0.46	0.458	0.452	0.45
Non-Zero Coeffs								
Cats (out of 459 / 57)	219	229	228	208	45	47	41	44
Demos (out of 15)		11		11		11		11
Extras (out of 5 / 4)			5	5			3	3

1.5 Regressions on Price Promotions

In order to test whether there was any evidence of plant-based meat competing with meat, I identified a random source of price variation in the data: region level price promotions (sales). In private conversations with the grocery chain, they explained that promotions are almost exclusively decided on by each of 11 regional pricing executives, with very little input from the corporate office. For small, less important UPCs, such as the PBM UPCs, these individual pricing executives are unlikely to consult any pricing models or other data, but instead often simply put certain items on sale by intuition, or for the reason that those items had not been on promotion in a long time.

In order to test this lack of coordination, I calculated the correlation of PBM products by region with other PBM products and all other meat products across all regions. While some larger UPCs (fresh cuts of chicken) had reasonable correlations across regions, nearly all PBM UPC correlations with themselves AND with other UPCs were unremarkable. In addition, I ran regressions of lagged quantity of PBM on price across each region individually, to see if some regions were putting PBM on promotion when either demand had been high or low in the weeks earlier, and found no significant results in any region. As a result, I felt confident in the assessment of my corporate contact at this chain that promotional prices for PBM are as good as random. This is especially true when looking at each specific store: decisions to promote are not related to store specific facts, as they are set at a higher regional level.

Promotional prices are identified in the data by two variables: net price and gross price. Gross price is the price listed on the price tag in the store as the price when the item is not on promotion (so the net price should equal the gross price). If the item is on promotion, the gross price is the price listed as "was", while net price is the price actually paid. In theory, all these prices are identical within each region each week, but sometimes there is unexplained variation in the prices that customers actually pay. This may be explained by manufacturer's coupons, but other times it is unexplained. In order to get a region level weekly promotion, I subtract the gross price from the net price across all purchases within a region in each week.

Figure 1.2: Prices of PBM Beef in Nor CA

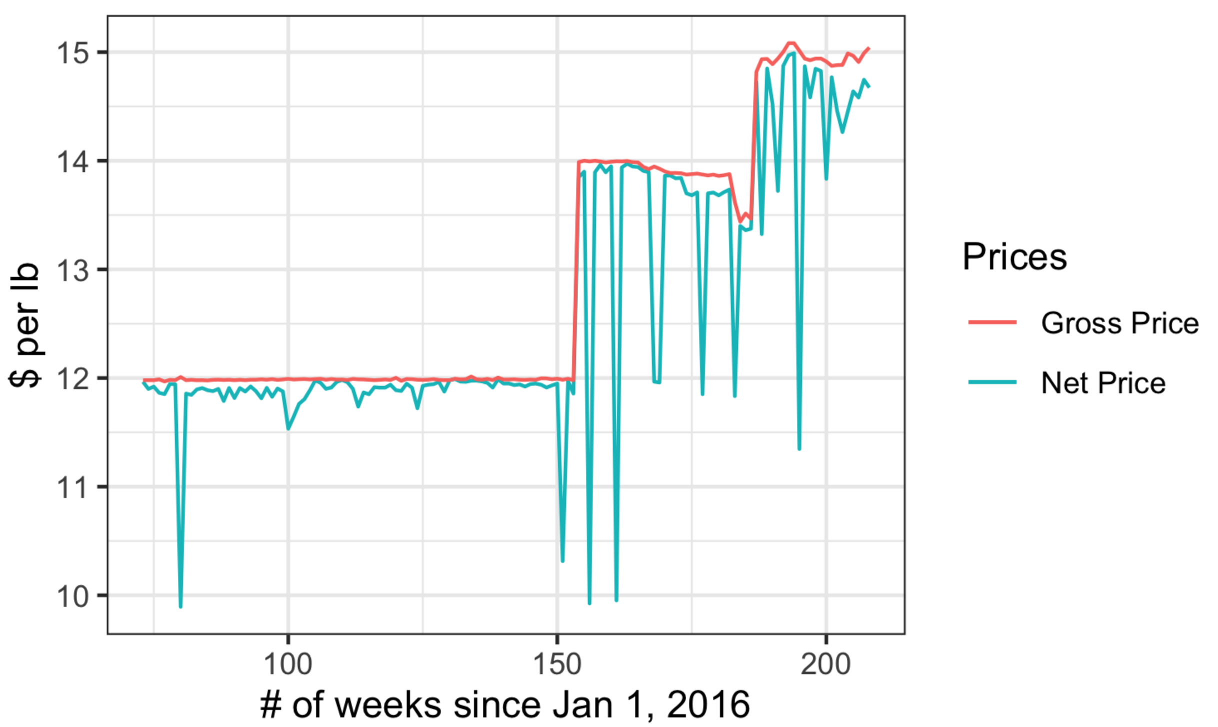


Figure 1.2 is an example of how promotions occurred in an average region over the study time. In some regions, semi-permanent promotions occurred, where the net price would be lowered for a long time, without the gross price being lowered. In these cases, after 4 weeks, I designated this as the new gross price and adjusted the gross price downward, so that no promotion could last longer than 4 weeks. Using this adjustment to gross price, I calculated a 'percent discount' as a plausibly random price shock. After summarizing each household's

purchases for each week (if they did not have a visit in a week, they were missing in that week), I then used percent discount to run the following regressions for PBM beef, ground beef, chicken, veggie burgers, and tofu:

$$D_{ntri} = \beta S_{tri} + \alpha p_{tr,j\neq i} + \delta H + \zeta W + \epsilon_{ntri}$$

Where D_{ntri} are the lbs bought that week by household n in week t in region r of good i. S_{tri} is the percent discount, $p_{tr,j\neq i}$ are the prices of the other goods $j \neq i$, H is a suite of household level fixed effects, and W is a suite of week level fixed effects. Standard errors were clustered at the region level. The important regressions for policy would regress PBM discount on demand for beef. However, I didn't even get that far, as the regression of PBM discount on PBM demand showed that they were unrelated in table 1.11:

Table 1.11: Regressing Promotional Prices on Demand

	Dependent variable:				
	PBM Beef	Ground Beef	Chicken	Veggie Burgers	Tofu
	(1)	(2)	(3)	(4)	(5)
Percent Discount	0.003	0.293^{***}	0.312^{***}	0.004^{***}	0.025^{***}
	(0.008)	(0.113)	(0.115)	(0.001)	(0.005)
Price PBM Beef		−0.00005	-0.005^{*}	−0.00002	0.0002
		(0.001)	(0.003)	(0.0001)	(0.0001)
Price Grd Beef	0.0003		−0.009	0.0001	0.0003
	(0.001)		(0.020)	(0.0002)	(0.001)
Price Chicken	0.0001	−0.004		−0.0002	−0.0001
	(0.001)	(0.008)		(0.0003)	(0.001)
Price Veg Burger	−0.0005	−0.001	−0.0005		−0.0001
	(0.001)	(0.003)	(0.006)		(0.0002)
Price Tofu	-0.003^{*}	-0.018^{***}	−0.012	-0.001^{*}	
	(0.002)	(0.005)	(0.017)	(0.0003)	
Observations	19,272,642	19,272,642	19,272,642	19,272,642	19,272,642
R^2	0.091	0.141	0.160	0.130	0.157
Adjusted R^2	0.081	0.132	0.151	0.121	0.148
Note:					*p<0.1; **p<0.05; ***p<0.01

At first, this was a bit mysterious, but the answer may have been supply shortages. During this time, PBM was brand new, and supply may have been limited. When it would go on promotion, it may (the data do not show this, it must be inferred) be that after a week, the stores would sell out. To test this, I categorized any percent discount of more than -0.08

as a promotion, and put the regressions from table 1.11 into an event time regression:

$$D_{ntri} = \beta E_{tri} + \alpha p_{tr,j\neq i} + \delta H + \zeta W + \epsilon_{ntri}$$

Where E_{tri} is weeks before or after a promotion for good i. the week before the promotion (week -1) is omitted. the value of β is displayed in figures 1.3. In subfigure 1.3a, we can see that while there was an impact on demand in the first week of a promotion, by the second week, there was no longer an effect. That seems to imply that supply issues may very well have been at fault. It's important to note that there were no effects on demand in the runup to the promotions, which supports the identifying assumption that the promotions were plausibly randomly timed. In order to see if these increases in PBM demand had any effect on demand for other goods, I altered the regression as follows:

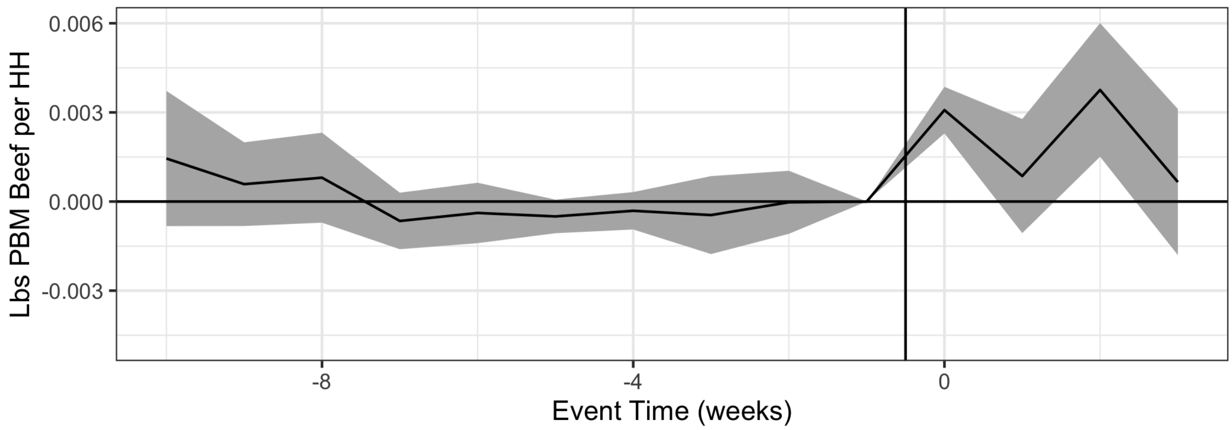
$$D_{ntri} = \beta E_{tr,PBM} + \alpha p_{tr,j\neq iorPBM} + \delta H + \zeta W + \epsilon_{ntri}$$

This way, the event time was still in weeks before and after a PBM promotion, and omitted PBM price. The results are presented in figures **??**. Those figures show that the promotions on PBM had no noticeable difference on the demand for ground beef, chicken, veggie burgers, or tofu. There are a number of reasons to think these regression results are not perfect. First, the number of clusters, 12, was quite low, but that was the level at which prices were set. Second, the predicted demand change for PBM beef in the first week of the promotion (0.003 lbs) was probably not large enough to overcome the normal noise in the much, much larger chicken and ground beef markets. Yet, the main finding remains that we are unable to detect any effect of PBM on the demand for meat.

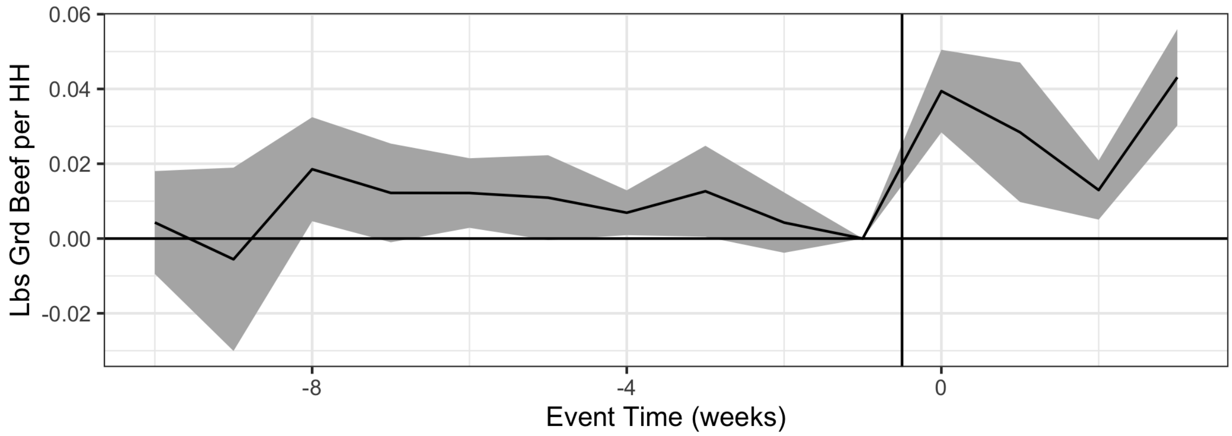
1.6 Conclusion

I examined by far the largest sample of PBM purchasers than has ever been available to researchers. I found that grocery store data, while offering plenty of information on the history of each customer, cannot explain PBM future purchases well. What can be explained is not terribly encouraging for those who are hoping that PBM will ultimately transition meat eaters away from meat. When I used plausibly randomly assigned PBM discounts, I was unable to find any movement away from meat. The market for PBM is very young, however. As more firms enter the market, prices will continue to fall and quality will improve. It may just take a lot longer to truly compete with meat than what some futurists and plant-based meat cheerleaders had predicted.

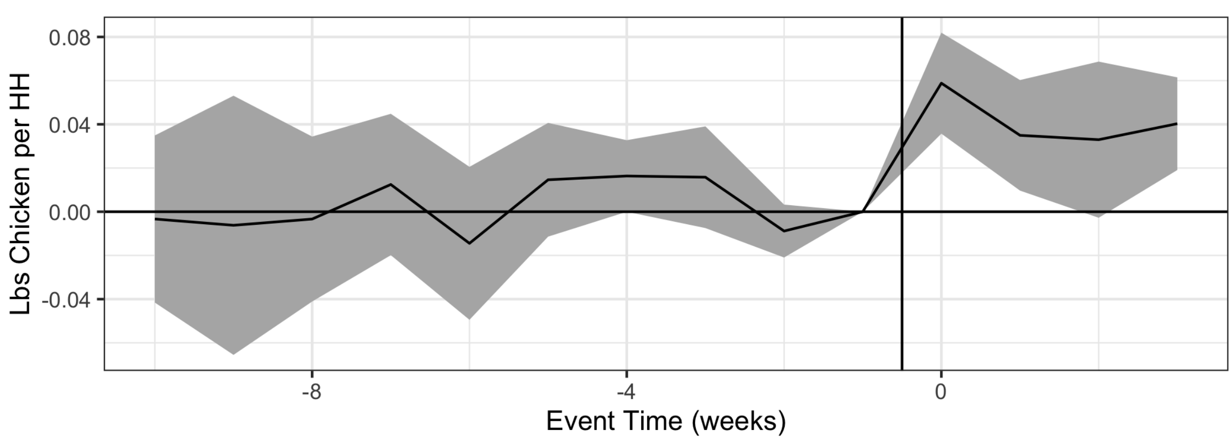
Figure 1.3: Event Study, impact of promotion on Lbs bought per Household



(a) PBM Beef Promotions

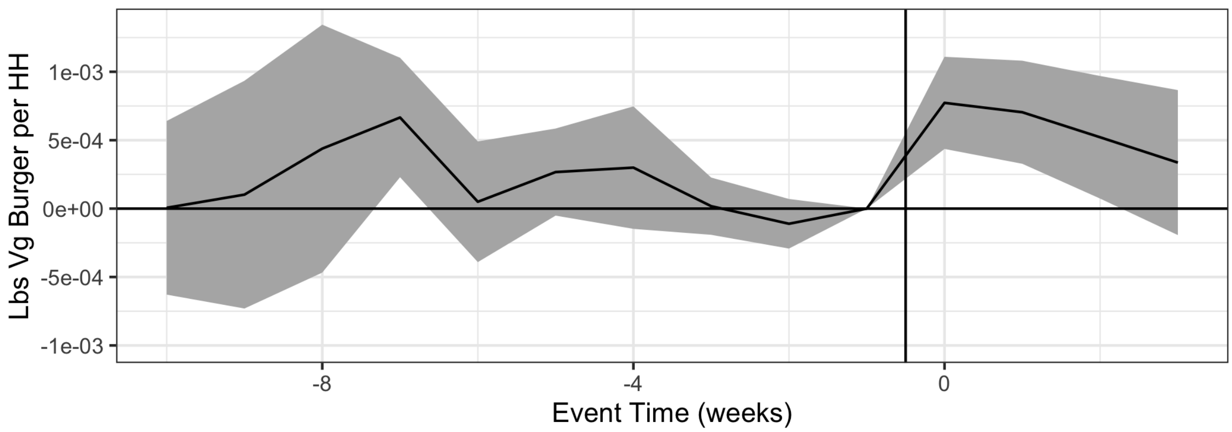


(b) Ground Beef Promotions

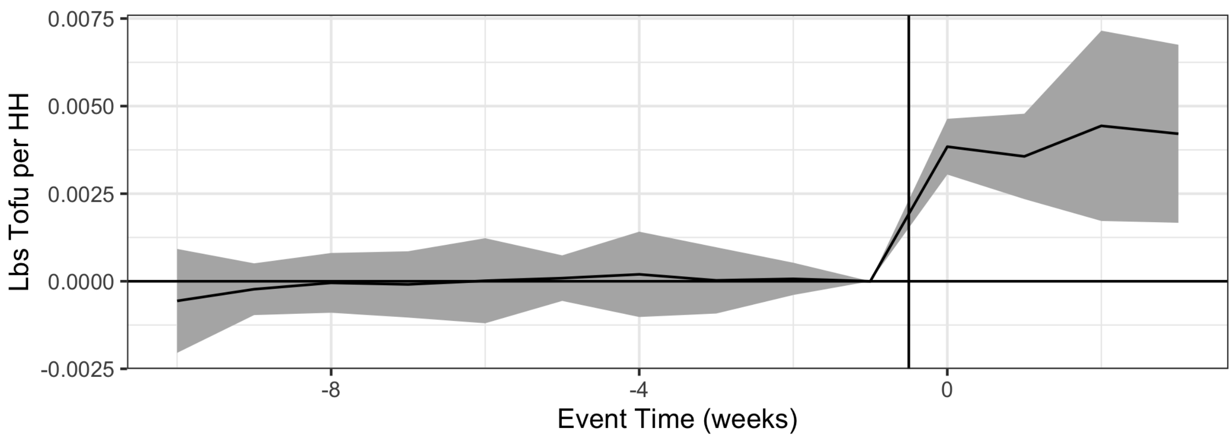


(c) Chicken Promotions

Figure 1.3: Event Study, impact of promotion on Lbs bought per Household(cont.)

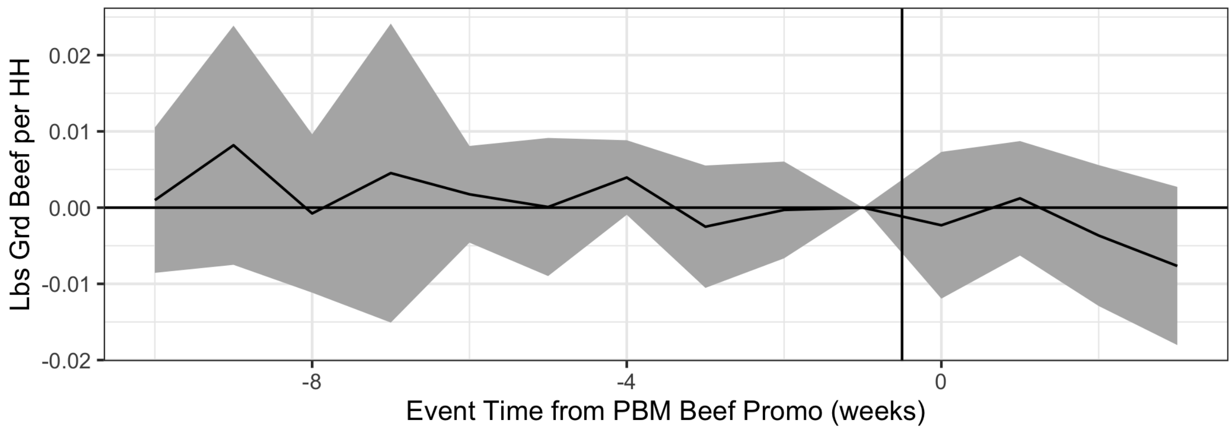


(d) Veggie Burger Promotions

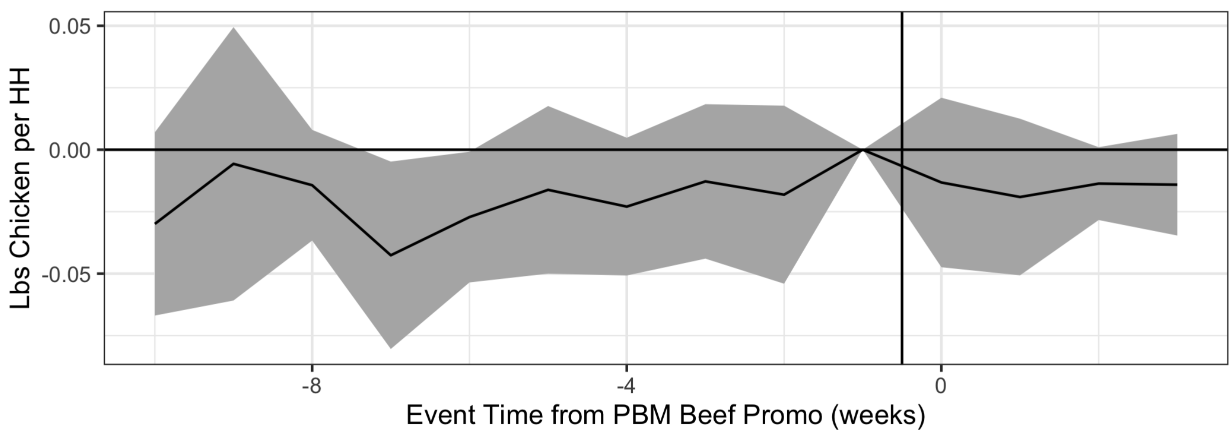


(e) Tofu Promotions

Figure 1.4: Event Study, impact of PBM Beef promotion on Lbs bought of OTHER items per Household

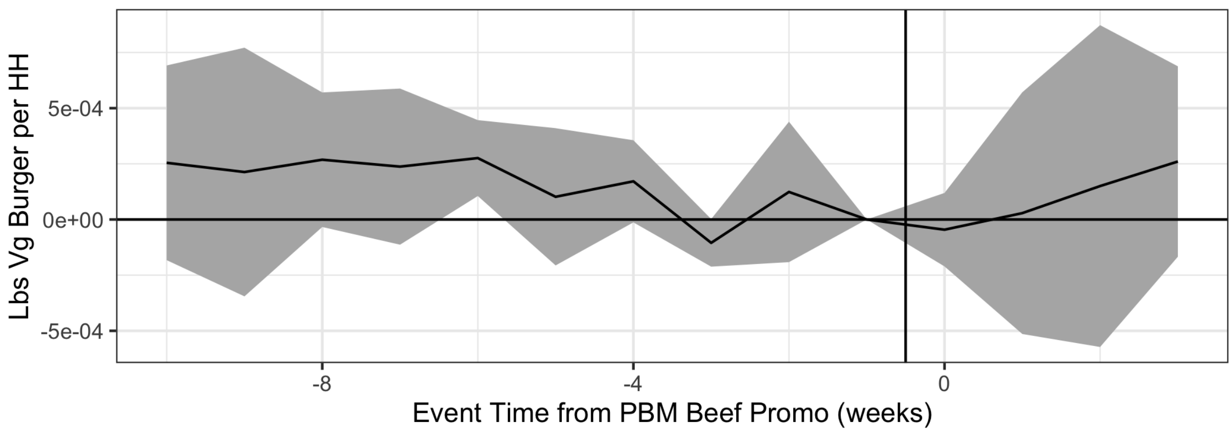


(a) Ground Beef

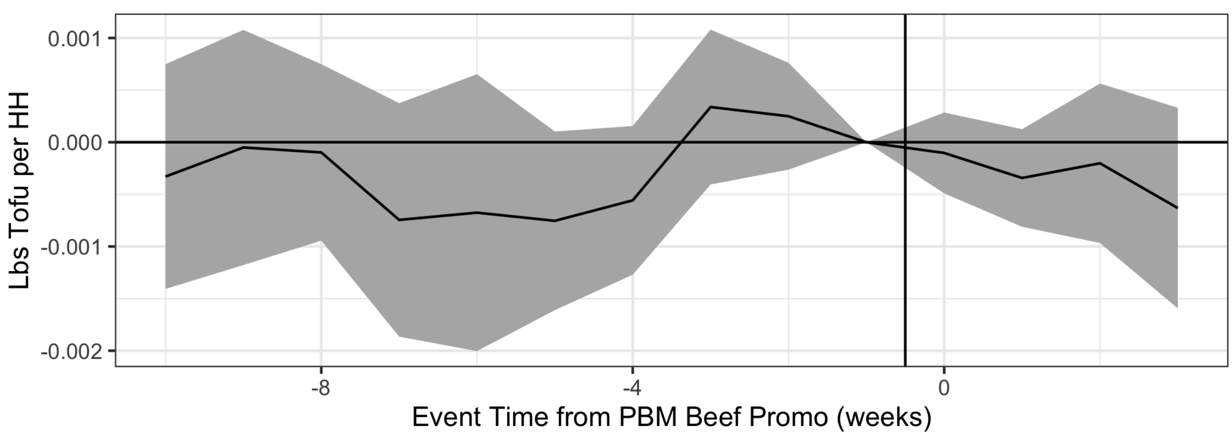


(b) Chicken

Figure 1.4: Event Study, impact of PBM Beef promotion on Lbs bought of OTHER items per Household(cont.)



(c) Veggie Burger



(d) Tofu

Bibliography

Carlsson, Fredrik, Mitesh Kataria, and Elina Lampi. 2022. "How much does it take? Willingness to switch to meat substitutes." *Ecological Economics* 193:107329.

Gerber, Pierre J, Henning Steinfeld, Benjamin Henderson, Anne Mottet, Carolyn Opio, Jeroen Dijkman, Allessandra Falcucci, Giuseppe Tempio, et al. 2013. *Tackling climate change through livestock: a global assessment of emissions and mitigation opportunities.* Food / Agriculture Organization of the United Nations (FAO).

Henders, Sabine, U Martin Persson, and Thomas Kastner. 2015. "Trading forests: land-use change and carbon emissions embodied in production and exports of forest-risk commodities." *Environmental Research Letters* 10 (12): 125012.

Huang, Lichun, Hal G Gordon, and David Zilberman. 2020. "The Challenges and Potential Benefits of Animal-Free Meats." *ARE Update* 23 (4): 5–8.

Mullainathan, Sendhil, and Jann Spiess. 2017. "Machine learning: an applied econometric approach." *Journal of Economic Perspectives* 31 (2): 87–106.

Neuhofer, Zachary T, and Jayson L Lusk. 2022. "Most plant-based meat alternative buyers also buy meat: an analysis of household demographics, habit formation, and buying behavior among meat alternative buyers." *Scientific Reports* 12 (1): 1–11.

Poore, Joseph, and Thomas Nemecek. 2018. "Reducing food's environmental impacts through producers and consumers." *Science* 360 (6392): 987–992.

Tibshirani, Robert. 1996. "Regression shrinkage and selection via the lasso." *Journal of the Royal Statistical Society: Series B (Methodological)* 58 (1): 267–288.

Upshot, New York TImes. 2018. "Political Bubbles and Hidden Diversity: Highlights From a Very Detailed Map of the 2016 Election." Accessed January 1, 2023. `https://www.nytimes.com/interactive/2018/07/25/upshot/precinct-map-highlights.html`.

Van Loo, Ellen J, Vincenzina Caputo, and Jayson L Lusk. 2020. "Consumer preferences for farm-raised meat, lab-grown meat, and plant-based meat alternatives: Does information or brand matter?" *Food Policy* 95:101931.

Zhao, Shuoli, Lingxiao Wang, Wuyang Hu, and Yuqing Zheng. 2022. "Meet the meatless: Demand for new generation plant-based meat alternatives." *Applied Economic Perspectives and Policy.*

Chapter 2

Recreational Damages from Air Pollution

2.1 Introduction

Note: this section was coauthored with Scott Kaplan
California experienced the worst wildfire season in it's history in 2018. The economic damages from these fires affected millions of people – tens of thousands lost their homes and other property, and millions of others experienced the terrible air quality that resulted from fires across the State. While there is substantial empirical research linking poor air quality to adverse health impacts, very little revealed preference empirical work has assessed the impact of poor environmental quality on demand for outdoor recreational activities – most of the empirical work has had to rely on non-market valuation methods because of the nature of these goods. Estimating damages from environmental externalities in different markets is extremely important for policy-making, and using a revealed preference approach allows us to do this in a rigorous manner.

This project leverages a unique setting to study the effects of air pollution on a market good people consume for recreational purposes: tickets to National Football League (NFL) games posted on a popular, online secondary marketplace. Specifically, we rely on substantial exogenous variation in air quality as a result of the severe California wildfires and implement a fixed effects model to estimate the causal impact of air pollution on prices of listed tickets. There were two games greatly impacted by the fires in California–the Oakland Raiders and San Francisco 49ers home games on November 11th and 12th, respectively. Both games were played as scheduled despite the terrible air quality, which was upwards of 150 AQI (classified as "Very Unhealthy") in the days leading up to and during the games. The sample used in our initial estimation contains these two games, as well as all other 49er and Raider home games that took place this past season.

Our initial findings suggest that an increase in the AQI does not lead to a statistically significant change in listed ticket prices (in fact, we find a slightly positive estimate that is likely the effect of unaccounted for noise). We also determine that there was no statistically significant impact on the number of tickets listed on the marketplace for these games. One may argue that this finding is the result of "lack of transparency" of adjustments in air quality on a day-to-day basis, not individuals' actual indifference between recreating under different air quality conditions. We feel confident ruling out this effect in our estimation, since the Camp Fire in Northern California caused the AQI in the Bay Area to increase by more than two standard deviations from the mean value found in our sample just before and during the two afformentioned games. Additionally, one might argue that there is a "lack of attention" on the secondary market to time-variant factors that may affect ticket prices. However, we do see the market respond to other sizeable and transparent events. For example, 49ers franchise quarterback Jimmy Garrapolo suffered a season-ending ACL tear on September 23rd. We find that this event led to a statistically significant 9.6% average reduction in listed prices for future San Francisco home games.

These results indicate that poor environmental quality may not detrimentally affect people's willingness-to-pay to recreate, at least in the case of National Football League games. Further research will implement this analysis in Major League Baseball (MLB) games (which we are currently collecting data for), which will act as a useful comparison since unlike football, there are many more games that are easily substitutable across one another if the environmental quality for a specific game turns out to be poor.

In addition to testing the impact of AQI on the demand for football games, we also examine if deviations from the historical average, game-time "feels-like" temperature (accounting for the heat index and wind chill) have an affect on the price fans are willing to pay for football tickets, and find tight zeros for these temperature deviation affects. Just like in our air quality analysis, we hope that the parallel analysis we plan to carry out on the effect of unexpected weather on the market for MLB games to see if consumers respond compared to the null response we find for football games.

The remainder of the paper will proceed as follows: we will conduct a brief overview of the literature to-date on the economics and valuation of consumer demand for recreation. We will then present the data used in this analysis as well as some key summary statistics. Next, we will present a simple conceptual framework guiding our empirical analysis of air quality on ticket prices. We will then present supplementary findings with respect to other events that may affect prices, including impacts on a team's competitiveness (i.e. through injuries to star players) and an initial analysis examining the impact of weather expectations on prices. Finally, we will conclude.

2.2 Literature Review

This paper fits at the intersection of the literature on the economic costs of air pollution and weather, valuation of recreation, and pricing in secondary ticket marketplaces. Environmental economists have been able to causally identify damage pathways from air quality via many pathways, including hospitalizations from asthma (Neidell (2009)), birth weight

(Currie and Walker (2011)), rates of dementia (Bishop, Ketcham, and Kuminoff (2018)), as well as non-health outcomes like productivity of agriculture workers (Graff Zivin and Neidell (2012)) and even MLB umpires (Archsmith, Heyes, and Saberian (2018)).

Finding air pollution impacts on measurable health or work outcomes is somewhat more straightforward than impacts on non-market goods like leisure. In fact, while there is a long literature measuring the valuation of outdoor leisure (Hanemann (1984)), especially when it comes to anglers (Egan et al. (2009)), we are unaware of any paper that seeks to measure the impact of air quality on recreation. There is some existing work on how temperature impacts leisure activity using the American Time Use Survey (Graff Zivin and Neidell (2014)) and bike share trips (Chan and Wichman (2018)). This paper contributes to the literature by identifying and studying a market for an outdoor recreation good that can be analyzed with revealed preference method.

High frequency markets for outdoor leisure are rare, so using the tickets market for professional sporting events affords a rare opportunity. Microdata from this market have already been used to test the theoretical models of dynamic pricing (Sweeting (2012)) and the economic value of star basketball players (Kaplan (2019)), but these data have yet to be used to measure the amenity value of air quality, which has important implications outside of the sports and ticket pricing world.

2.3 Data and Summary Statistics

To test the impact of environmental quality on willingness-to-pay to attend NFL games, we utilize high temporal frequency data on listings from a popular online secondary ticket marketplace, as well as hourly data on the air quality index in San Francisco and Oakland. More specifically, we use novel data-scraping methods to collect (i) individual ticket listing level-data for each future NFL game each half-hour from this online secondary marketplace, and (ii) hourly air quality data (and forecasts) from AirNow.gov in cities with NFL teams.

Secondary Ticket Marketplace Data

Table 2.1 exhibits summary statistics at the listing-level for all home San Francisco 49ers and Oakland Raiders games from the 2018 regular season. One can see that we have over 60 million listing by "Scrape ID" observations (data was scraped each half hour) across all of these matchups. The average listed price for a ticket across all scrapes was \$182.31, and the average number of tickets per listing was just over 3.

Table 2.2 provides additional summary statistics – in this case grouped at the matchup level. Therefore, the "Number of Observations" column refers to the mean, standard deviation, minimum, and maximum number of listing by half-hours-to-game observations present for *each matchup*, respectively. The "Average Listing Price" and the "Average Quantity per Listing" columns represent summary statistics on the matchup-specific averages. Therefore, \$187.30 is the average of the average listing price of each matchup. Finally, the "Number

Table 2.1: Summary Statistics (All Games)

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Listing Price	58,557,191	180.32	230.52	6.99	96.00	209.80	10,998.90
Quantity per Listing	58,557,191	3.18	2.46	1	2	4	37

of Unique Listings," "Number of Unique Scrapes," and "Number of Unique Zones" columns represent the mean, standard deviation, minimum, and maximum of the matchup-specific counts of these variables. "Zones" represent specific areas of each of the teams' stadiums. One can see that we observe an average of just over 5,000 unique listings for each matchup, providing substantial variation in listed prices and quantities.

Table 2.2: Summary Statistics (Grouped by Matchup)

	Number of Observations	Average Listing Price ($)	Average Quantity per Listing	Number of Unique Listings	Number of Unique Scrapes	Number of Unique Zones
Mean	4182656.50	186.50	3.20	5279.71	2684.00	17.21
SD	2980762.21	63.35	0.33	1981.11	1490.00	0.97
Min	196521.00	137.04	2.62	2302.00	176.00	16.00
Max	9184347.00	354.87	3.77	8468.00	4622.00	19.00

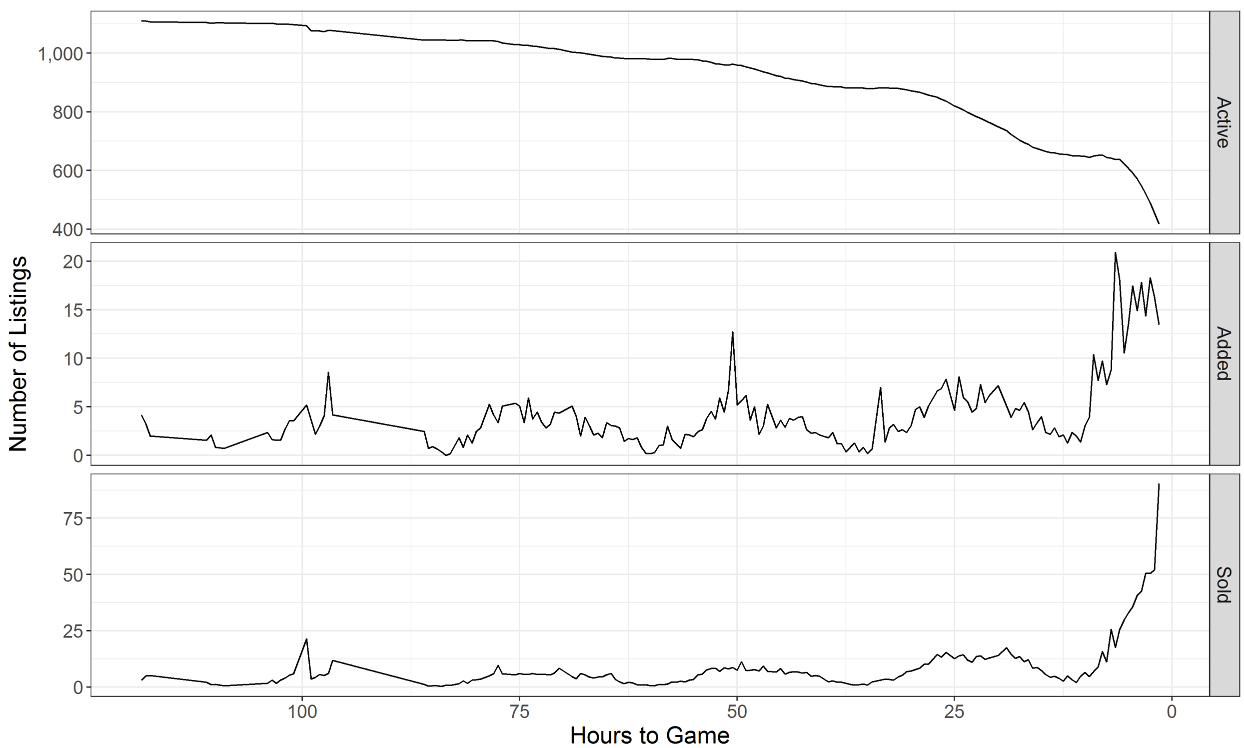
Finally, Figure 2.1 displays the average number of "Active," "Added," and "Sold" unique listings by hours to game across all San Francisco 49ers and Oakland Raiders home games. "Active" ticket listings refer to those that are available on the marketplace at a given hours-to-game value. The average number of "Added" and "Sold" listings refer to the number of unique listings added and sold, respectively, compared to the previous hours-to-game value. There are a couple of notable trends to point out. First, the average number of active listings on the marketplace declines as a matchup approaches. This is because more listings are sold than added over this timeframe.[1] Next, there is a more uniform pattern of listings added to the marketplace before a matchup, while the number of sold listings spikes in the several hours just before a matchup. These trends are quite consistent across different types of matchups (i.e. in terms of competitiveness, time of day or week, etc.).

Air Quality Data

The second set of data we utilize is the hourly observed PM2.5 air quality index for each of the two cities where these teams play, collected from AirNow.gov. PM2.5 consists of

[1] Note the different y-axis scales in each of the three panes of Figure 2.1.

Figure 2.1: Average Number of Listings Active, Added, and Sold by Hours to Game



particulates resulting from fires or other localized pollution events and are especially harmful to individuals if inhaled. The index is a measure of how many PM2.5 particles are in the air during each hour of measurement, where a value of “100" represents the benchmark for a city to be “in compliance" with air quality standards. To better understand the relationship between the AQI and health concerns, Table 2.3 below shows the AQI value ranges and their corresponding categories of quality. The index takes values between 0-500, ranging from good to hazardous air quality.

Figure 2.2 shows the air quality in each of these regions during the month of November, when the fires had led to drastic AQI increases. Our identifying variation comes from the air quality spikes occurring before the November 11th Oakland Raiders game and the November 12th San Francisco 49ers game. One can see the average AQI nearly triples to just under 200 in the few days before each of these games, which falls in the “Unhealthy" category of the AQI range, where everyone may begin to experience health effects.[2] Our expectation was that this reduced environmental quality would have led to lower equilibrium prices on the secondary ticket marketplace, may have increased the number of tickets listed on the

[2]The AQI actually gets worse after these games, but both of these teams did not play home games affected by these higher air quality values.

Table 2.3: Air Quality Index Categories and Health Descriptions

AQI Index Range	Category	Description
0-50	Good	Air quality is considered satisfactory, and air pollution poses little or no risk.
51-100	Moderate	Air quality is acceptable; however for some pollutants there may be a moderate health concern for a very small number of people who are unusually sensitive to air pollution.
101-150	Unhealthy for Sensitive Groups	Members of sensitive groups may experience health effects. The general public is not likely to be affected.
151-200	Unhealthy	Everyone may begin to experience health effects; members of sensitive groups may experience more serious health effects.
201-300	Very Unhealthy	Health warnings of emergency conditions. The entire population is more likely to be affected.
301-500	Hazardous	Health alert; everyone may experience more serious health effects.

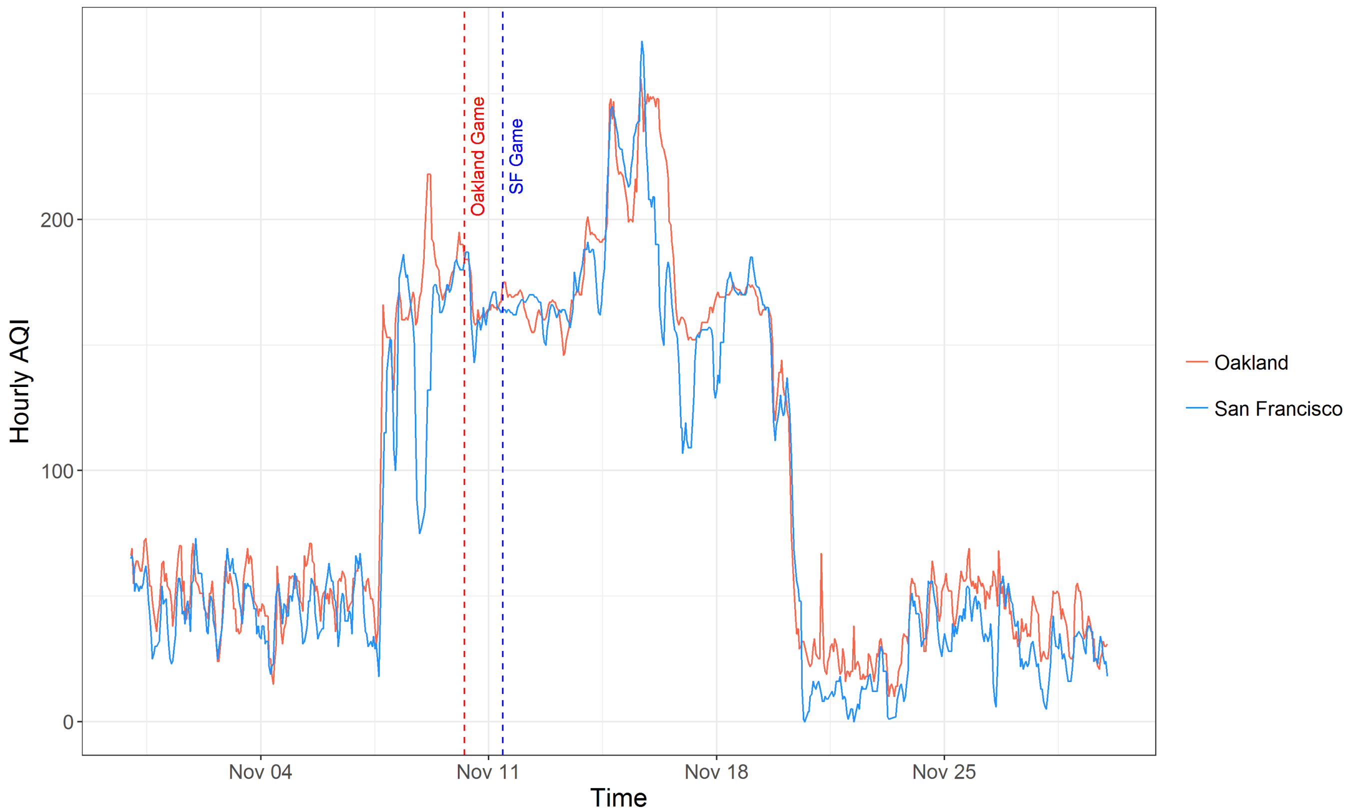
Source: First Environments Early Learning Center. http://www.firstenvironments.org/aqi

secondary marketplace since, or both, considering that football games in these stadiums are played outdoors and last several hours.

Descriptive Effect of Poor Air Quality on Quantities and Prices

We are interested in examining the impact of the poor air quality event caused by the Northern California wildfires on willingness-to-pay to attend affected NFL games. Figure 2.3 presents the average price trajectories of listed tickets by hours-to-game for each Oakland Raiders (left pane) and San Francisco 49ers (right pane) home game. The treated games are shown in red, while the non-treated games are shown in blue. First, one can notice that prices decline (on average) in the hours leading up to a game. Additionally, price volatility also increases in the several hours before a game. However, one can also see that there is no discernible price effect in the treated games associated with the AQI event compared to the

Figure 2.2: Hourly Air Quality Index for Oakland and San Francisco around Impacted Games



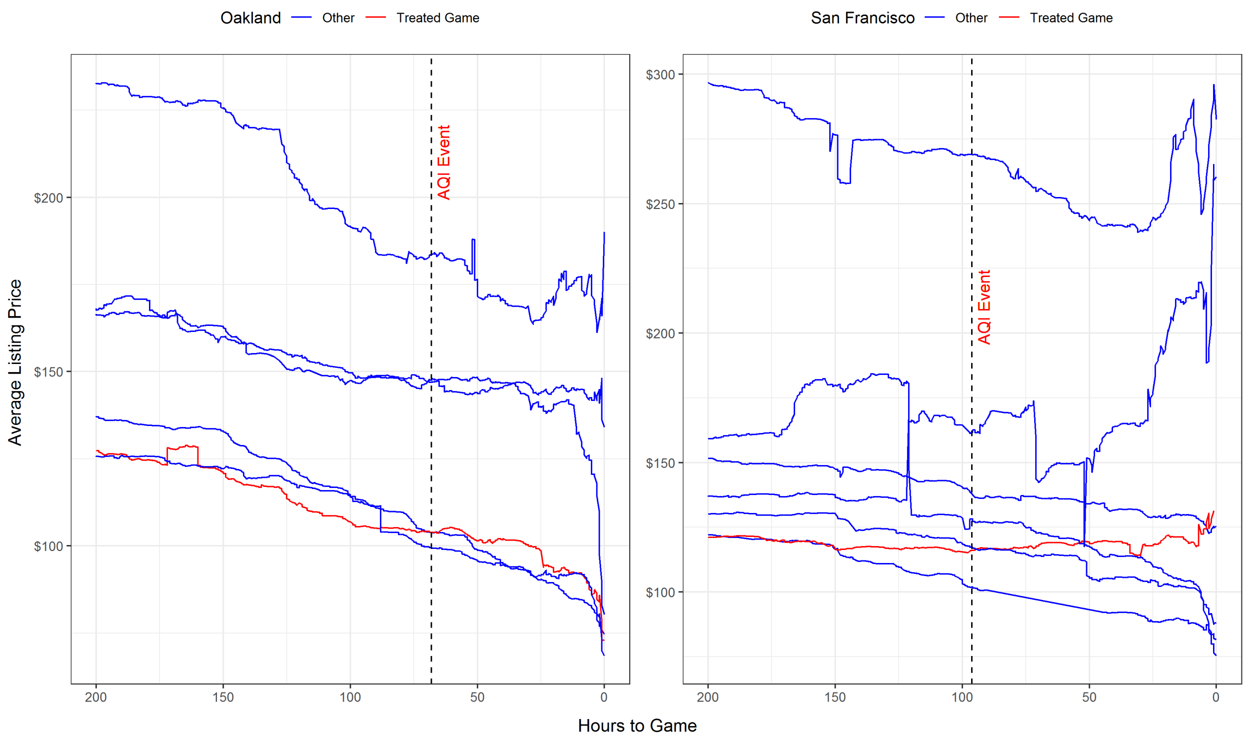
control games.[3] This provides initial evidence that people may not have responded to this poor air air quality event in terms of their willingness to pay.

It is also important to determine the effect of the air quality event on quantity of tickets listed on the marketplace for the treated games. Our prior is that quantity should be unadjusted in response to an exogenous shock, assuming that ticket sellers and buyers respond uniformly (supply increases while demand decreases, but by the same amount), leading to a decrease in equilibrium price and an unchanged equilibrium quantity. However, it is plausible that many people may have listed their tickets at the going price on the marketplace at the time of this air quality event, and prices may not have adjusted at that time since people were using already listed prices as the benchmark for determining their prices. This would indicate some type of negative demand response to the air quality event.

Figure 2.4 displays the number of active listings per game for Oakland (left-pane) and

[3] We define the "AQI event" as the point in time when the AQI spiked from below 50 in each city to above 150 in the span of 3 hours on the afternoon of November 8th.

Figure 2.3: Average Price of Tickets for all Bay Area NFL Games



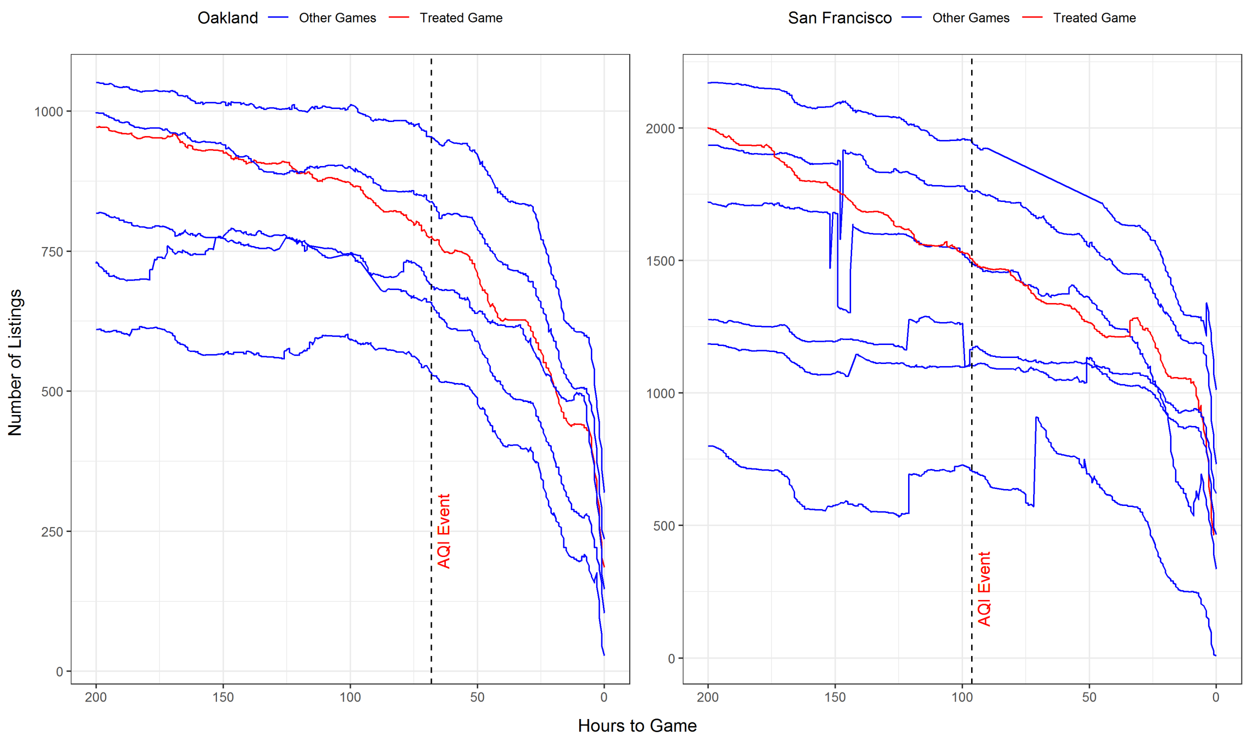
San Francisco (right-pane) home games. Again, treated games are shown in red while control games are in blue. One can see that there are no differential changes in quantity in the treated games as a result of the air quality event. Additionally, Figure 2.5 showcases the number of active, added, and sold listings by hours-to-game for each of the two treated games, where the blue and red dotted lines indicate the time of the AQI event in terms of hours-to-game for the treated San Francisco and Oakland games, respectively. Again, there doesn't seem to be any discernable changes in the number of listings added or sold in response to these events.

These figures provide descriptive evidence of null price and quantity responses to the negative AQI event. In section 2.4, we will present an empirical framework that will be used to determine a causal estimate of AQI changes on listed prices and quantities on the secondary marketplace. Section 2.5 will then present the results of this estimation.

2.4 Conceptual Framework

The primary empirical strategy for this analysis relies on identifying the causal impact of changes in air quality on willingness-to-pay to attend NFL games. This will be achieved

Figure 2.4: Number of Listings Available for all Bay Area NFL Games



using a fixed effects model, where the identifying variation comes from changes in air quality that is uncorrelated with other unobservables that may impact ticket prices or quantity of tickets listed. The estimating equation for price will be as follows:

$$Price_{lsih} = \beta_1 AQI_{ih} + \alpha_h + \alpha_{is} + \epsilon_{lsih} \tag{2.1}$$

where $Price_{lsih}$ represents the price of listing l in stadium section s for matchup i at hours-to-game h. AQI_{ih} is the PM2.5 air quality index for matchup i at hours-to-game h. Finally, α_h and α_{is} represent hours-to-game and section-by-matchup fixed effects, respectively.

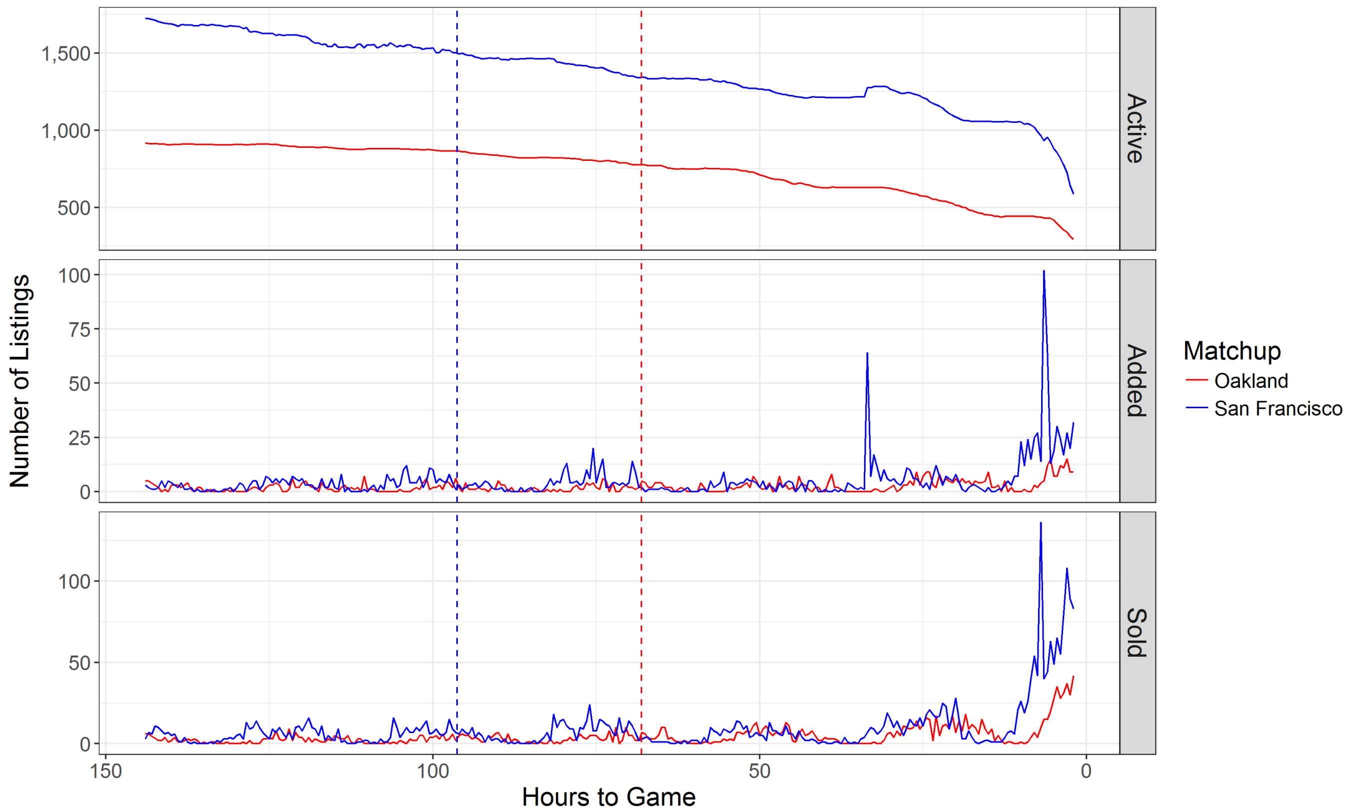
We also estimate equation 2.1 with the quantity of listed tickets as the dependent variable. We write this specification as follows:

$$Quantity_{sih} = \beta_1 AQI_{ih} + \alpha_h + \alpha_{is} + \epsilon_{sih} \tag{2.2}$$

where each quantity observation represents the total number of tickets found in section s for matchup i at hours-to-game h.

Willingness-to-pay to avoid a marginal increase in AQI is identified through the combination of equations 2.1 and 2.2. Namely, if price adjustments are negative and statistically

Figure 2.5: Total Number of Listings Active, Added, and Sold by Hours to Game for Treated Games



significant, and quantity adjustments are not significantly different from zero, then our model of ticket supply and demand suggests that price adjustments can be interpreted as a change in willingness-to-pay.

Threats to Identification

Our primary identifying assumption is that:

$$\mathbb{E}[\epsilon_{lsih}|AQI_{ih}, \alpha_h, \alpha_{is}] = 0 \tag{2.3}$$

Namely, that we satisfy strict exogeneity when estimating β_1, which is the impact of AQI adjustments on the listed price of tickets. However, any time-varying unobservables correlated with AQI would bias our estimate of β_1.

The most likely omitted variable that may bias our estimate of β_1 is other weather-related factors that might affect ticket prices or quantities. For example, if high temperatures are correlated with poor air quality, and high temperatures also negatively impact ticket prices, then our estimate of β_1 would be downward biased (more negative) since it will attribute temperature-related impacts on prices to increases in AQI. However, because there is relatively little day-to-day variation in AQI, with the exception of the wildfire event, it is unlikely that a meaningful temperature event would be correlated with the large spike in AQI associated with the Northern California wildfires. Thus, we believe the meaningful day-to-day AQI variation in our sample is quite exogenous, and should provide a plausibly causal estimate of the change in prices or quantities as a result.

2.5 Results

Here, we present the results of our primary empirical estimation: the effect of air quality changes on listed prices and quantities of tickets on the secondary marketplace. Table 2.4 presents four separate specifications based on equation 2.1. Columns (1) and (3) use log($Price$) as the dependent variable, while columns (2) and (4) use $Price$ as the dependent variabe. Additionally, columns (1) and (2) use all listings in the estimation, while columns (3) and (4) use only listings that were *sold* at some point.[4]

One can see from Table 2.4 that the estimate of the effect of a 1-unit increase in AQI on listed prices is positive and statistically significant. This is counter-intuitive – one would think that increases in the AQI would lead to lower listed prices. Since a 1-unit change in AQI is not very intuitive, we translate these results into a 1 standard deviation change in AQI (first row) and the actual AQI increase as a result of the wildfires (second row). We find that a 1 SD increase in AQI led to an average ticket price increase of %2.38, which translates to $3.36. For additional context, the Northern California wildfires led to a 2.8 SD increase in AQI prior to the affected NFL matchups. These results are consolidated in Table 2.6. It is likely that the positive and significant nature of these estimates is due to noise in the data, but it is clear that changes in AQI do not seem to have a statistically significant *negative* impact on listed prices for NFL games.

Table 2.5 provides values from the estimation of equation 2.2. Again, columns (1) and (3) use log($Price$) as the dependent variable, while columns (2) and (4) use $Price$ as the dependent variabe, and columns (1) and (2) use all listings in the estimation, while columns (3) and (4) use only listings that were *sold* at some point. Here, we find that the quantity of tickets listed on the secondary marketplace does not respond to changes in the AQI, as our estimates are not statistically different from zero. The combination of Tables 2.4 and 2.5 suggest that there is a null price and quantity response to AQI changes, which is supported by the graphical evidence presented in section 2.3.

One may argue that these findings are the result of “lack of transparency" of adjustments in air quality on a day-to-day basis, not individuals’ actual indifference between recreating under different air quality conditions. We feel confident ruling this out in our estimation, since the Camp Fire in Northern California caused the AQI in the Bay Area to increase

[4]We assume a listing is sold if it disappears at a certain hours-to-game before the matchup.

by nearly three standard deviations from the mean value found in our sample just before and during the two aforementioned games. Additionally, one might argue that there is a “lack of attention" on the secondary market to time-variant factors that may affect ticket prices. However, as presented in the next section, the market responds to other sizable and transparent events. For example, 49ers franchise quarterback Jimmy Garoppolo suffered a season-ending ACL tear on September 23rd. We find that this event led to a statistically significant 9.6% reduction in listed prices for future San Francisco home games. We also find that in over half of all listings, prices are adjusted by sellers at least once, suggesting that a sizeable portion of the market is relatively active in adjusting prices of their listed tickets.

2.6 Evidence of Other Impacts on Ticket Prices

One potential objection to this research design is that by measuring the changes in posted prices, we are unlikely to see any signal from changes in valuation if there is a reasonable amount of rational inattention. While it may be that impact of air quality has such a small impact on the value of outdoor leisure that few ticket sellers would bother adjusting, we do find plenty of evidence that sellers are active in this market. Sweeting 2012 found that ticket sellers follow a predictable dynamic pricing schedule, keeping ticket prices high and then rapidly dropping them as the event approaches. We find a similar pattern in our data, and in Figure 2.6, we find that sellers change the price of their listings at least once more than half the time, and more than 10% of listings have very attentive sellers, with at least 11 price changes before they are either sold or the game starts. Even if sellers are not reacting to the air quality event specifically, but dynamically to demand, if they are attentive enough and demand is impacted by the air quality event, we would be able to see a drop in value by game-time.

In fact, we find a large impact in ticket value from another example of an unexpected event that changed the quality of attending NFL games in the Bay Area. Figure 2.7 uses an event-study framework to display the average impact on ticket prices as a result of an announcement that the 49ers star quarterback, Jimmy Garoppolo, would be out for the season with a torn ACL. The estimating equation used for this analysis is:

$$\log(Price_{lsiht}) = \sum_{t=-7}^{7} \boldsymbol{D}_t \boldsymbol{Announcement}_{t,ih} + \alpha_h + \alpha_{is} + \epsilon_{lsiht} \tag{2.4}$$

where t represents event-time in the context of scrapes that ocurred before and after the injury announcement. For example, $t = 1$ refers to all scrapes collected for all future San Francisco 49ers games taken the day after the announcement. We omit $t = -1$ for collinearity purposes, and bin the endpoints of this estimation at $t = [-7, 7]$, where the left and right endpoints represent the average of all periods before and after, respectively. $\boldsymbol{D}_t$ represents the vector of treatment coefficients in event-time associated with the announcement. As can be seen in the figure, parallel pre-trends in listing prices for future 49ers games is satisfied,

Table 2.4: Effect of Air Quality on Secondary Ticket Prices

	Dependent Variable: Listed Price			
	log(Price)/All Listings	Price/All Listings	log(Price)/Sold Listings	Price/Sold Listings
AQI	0.00228***	0.06208**	0.00056**	0.07356**
	(0.00000)	(0.03069)	(0.00025)	(0.03534)
Matchup*Section FE	Yes	Yes	Yes	Yes
Hours to Game FE	Yes	Yes	Yes	Yes
Clustered Robust SEs (Matchup-Level)	Yes	Yes	Yes	Yes
Observations	2,488,654	2,488,654	1,898,148	1,898,148
R^2	0.70020	0.14808	0.71276	0.13316
Adjusted R^2	0.69985	0.14791	0.71269	0.13295

Note: *p<0.1; **p<0.05; ***p<0.01

Table 2.5: Effect of Air Quality on Quantity of Secondary Tickets Listed

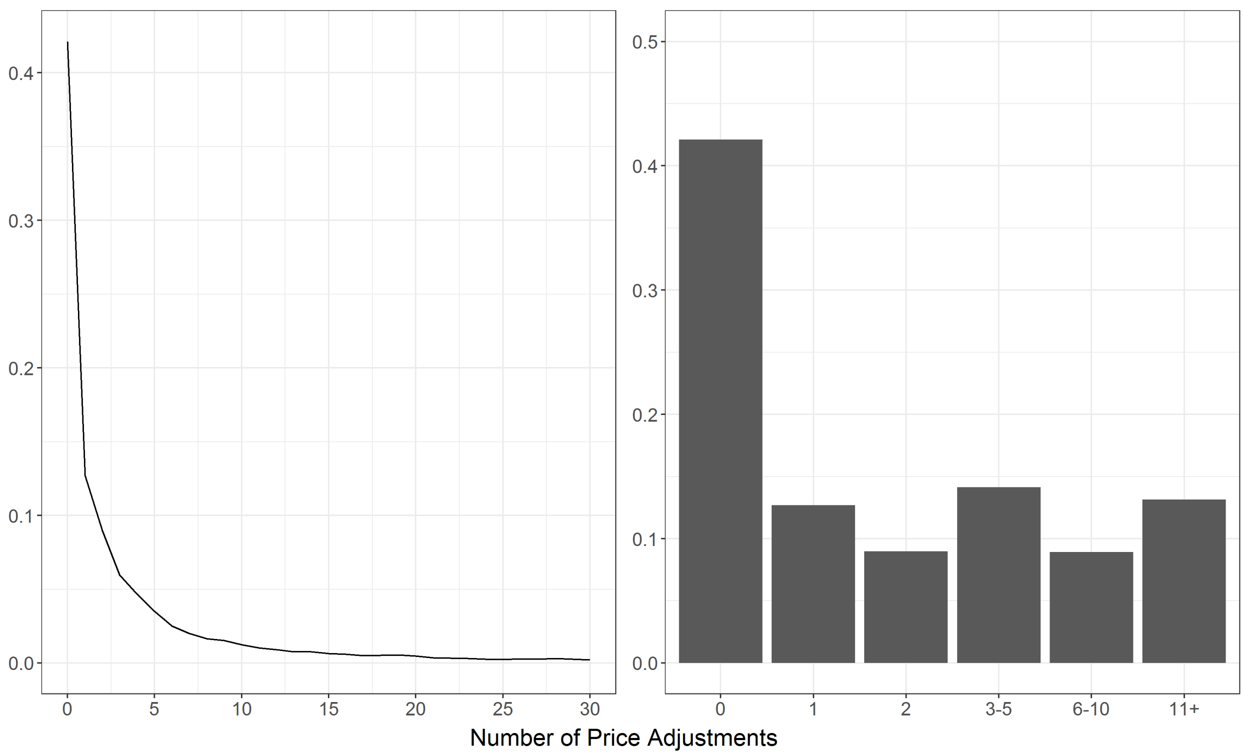
	Dependent Variable: Quantity of Tickets			
	log(Quant)/All Listings	Quant/All Listings	log(Quant)/Sold Listings	Quant/Sold Listings
AQI	−0.00053	−0.02272	0.00023	0.00373
	(0.00044)	(0.09269)	(0.00057)	(0.07652)
Matchup*Section FE	Yes	Yes	Yes	Yes
Hours to Game FE	Yes	Yes	Yes	Yes
Clustered Robust SEs (Matchup-Level)	Yes	Yes	Yes	Yes
Observations	2,488,654	2,488,654	1,898,148	1,898,148
R^2	0.97113	0.95299	0.96597	0.91825
Adjusted R^2	0.97112	0.95298	0.96597	0.91823

Note: *p<0.1; **p<0.05; ***p<0.01

Table 2.6: Interpretation of Price Estimation Results

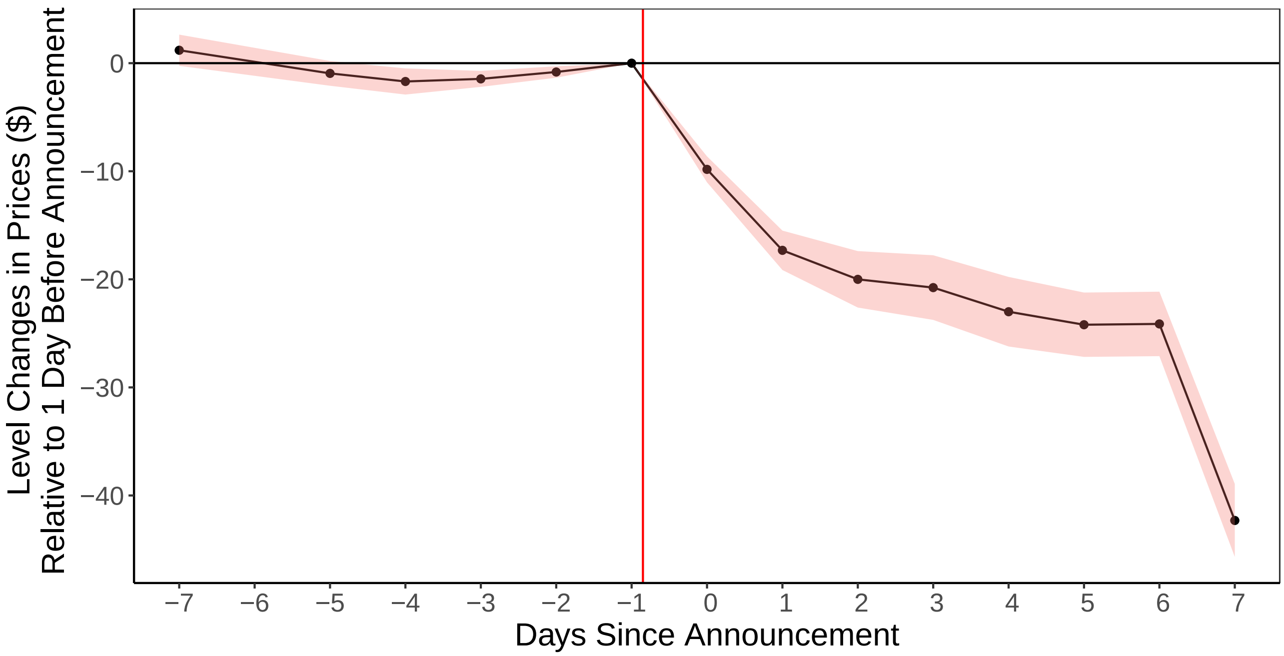
Scenario	AQI Change	Percent Increase in Average Ticket Price	$ Increase in Average Ticket Price
1 SD Increase	42.18	2.38%	$3.36
Wildfire-Related Increase	118.31	6.67%	$9.42

Figure 2.6: Density of Ticket Listings by Number of Times Price was Adjusted



since the values of $D_{-7}, ..., D_{-1}$ are nearly statistically indistinguishable from zero. However, for scrapes beginning the day of the injury announcement corresponding to listings for all future 49ers games, prices begin to fall quite substantially. Similar to equation 2.1, we control for matchup by section fixed effects, as well as hours-to-game fixed effects.

Figure 2.7: Event Study – Impact of J. Garoppolo Injury Announcement on Ticket Prices



We also present the results of a classic Difference-in-Differences (DID) estimation of prices in response to the Jimmy Garoppolo injury announcement in Table 2.7, using Oakland Raiders home games as the control group. The estimating equation for this analysis is:

$$\log(Price_{lsihz}) = \beta_1(PostAnnouncement * SFGames)_{iz} + \alpha_h + \alpha_{is} + \epsilon_{lsihz} \tag{2.5}$$

where z represents the actual scrape time ticket listing data was collected. Thus, for scrapes taken on or after September 24, 2018, $PostAnnouncement_z = 1$ (the date of the injury announcement). $SFGames_i = 1$ for all San Francisco 49ers games. Thus, β_1 represents the DID treatment coefficient, interpreted as the causal impact of the announcement on ticket prices. We compare the estimates from this DID estimation to our primary specification for estimating the price response to AQI changes. Columns (1) and (2) present our primary specifications from the AQI analysis, which correspond to columns (3) and (4) of Table 2.4, and columns (3) and (4) the corresponding DID estimations for the Jimmy Garoppolo injury. Again, one can see that the Jimmy Garoppolo estimation shows large and statistically significant decreases in listed prices of tickets following the injury announcement, while there seems to be a positive impact of AQI increases on listed prices (which we hypothesized earlier to be the result of noise in our data – we interpret these estimates as null effects).

Table 2.7: Comparing Price Impacts from Air Quality to Jimmy Garoppolo Injury Announcement

	Dependent Variable: Listed Prices (Sold Listings)			
	log(Price)	Price	log(Price)	Price
AQI	0.00056**	0.07356**		
	(0.00025)	(0.03534)		
Garoppolo Treat * Garoppolo Post			−0.10136***	−36.22838***
			(0.02616)	(9.91274)
Matchup*Section FE	Yes	Yes	Yes	Yes
Hours to Game FE	Yes	Yes	Yes	Yes
Clustered Robust SEs (Matchup-Level)	Yes	Yes	Yes	Yes
Observations	1,898,148	1,898,148	56,982,001	56,982,001
R^2	0.71276	0.13316	0.70061	0.02781
Adjusted R^2	0.71269	0.13295	0.70058	0.02772
Note:			*p<0.1; **p<0.05; ***p<0.01	

2.7 Preliminary Weather Impact Analysis

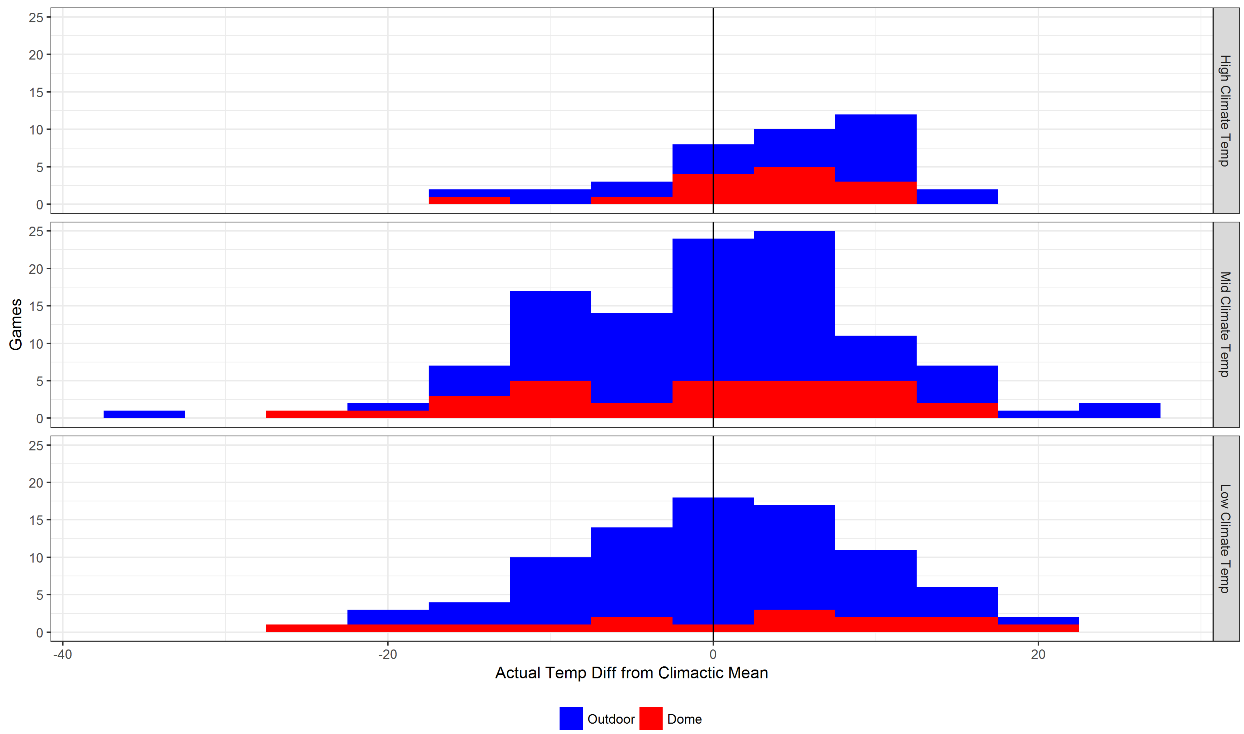
In addition to examining the effect of air quality, we also examined effect of temperature deviations from the climactic mean. We used data from the Dark Sky API, a weather prediction service that uses interpolation to provide super local weather predictions and historical realizations.[5] Using this data source, we collected the climactic average and the realized "feels-like" temperature (temperature plus heat index or wind chill) at game-time for all NFL regular season games in the 3rd week (the week that we had at least a weeks worth of scrapes prior to game-time) to the end of the season (208 days). Because fans are likely to respond differently to unexpectedly warm weather when the climactic average is warm versus cold, we split the sample into three groups: games whose climatic game-time average was above 70°F, those games where the climactic average was below 50°F, and those games in the middle. Figure 2.8 displays the spread of the difference in realized weather from the climactic average in each group.

The causal identification in this exercise is derived from these deviations in weather. From our scrapes, we sampled all listings from each game between 4 and 5 days (96 and 120 hours) to game-time and within 1 day of game-time.[6] The weather report 5 days out from a game is a very noisy predictor of temperature, but the weather report one day out is much more accurate. Thus, we make the assumption that buyers and sellers assume the actual

[5] https://darksky.net/dev

[6] As a robustness check, listings within 6 hours of the game were excluded, as occasionally the price can fluctuate unpredictably right before game-time, but this had no appreciable affect on our findings.

Figure 2.8: Feels-like Temperature Deviation from Climactic Average



weather will be close to the climactic mean 5 days out, but one day out, assume it will be close to the actualized weather.[7]

Because our sample includes a large number of games played in domes (where outdoor temperature should have little effect on the value of the tickets), we are able to use a triple difference design:

$$Price_{lzih} = \beta_1 D_h + \beta_2 T_i * D_h + \beta_3 T_i * O_i + \beta_4 T_i * O_i * D_h + \alpha_i + \alpha_z + \epsilon_{lzih} \tag{2.6}$$

Where D_h is a dummy for the hour h of scrape being within one day of game-time, T_i is the temperature deviation for game i, O_i is a dummy for game i being played in an outdoor stadium, α_i is a fixed effect for game i and α_z is a fixed effect for the seating zone of listing l. The triple-difference coefficient of interest is β_4. Equation 2.6 was run separately for each climate group, allowing β_4 to have a different sign for warm and cold games, and standard errors were clustered at the game level.

[7]Picking 5 days also had the advantage that no games would be played in the intervening time, meaning fans had no better idea of the quality of the teams of playoff implications 5 days out versus one day out. We also tried the regressions comparing ten days out to 1 day out with no appreciable difference in results. For our future work using MLB ticket sales, we are conducting real time scrapes of weather predictions.

The results of the regression are displayed in Table 2.8. Unsurprisingly, we find that the price of the average listing is much lower a day out than 4 to 5 days out. This confirms previous literature about ticket pricing dynamics. It also appears that the effect of being one day out difference by stadium type, although this may be picking up something about the specific markets of the handful of domed stadiums, and is unlikely to be causal. The triple difference, importantly is a tight zero for all three climates, meaning there does not appear to be an appreciable effect of unexpected temperature on the value of attending an NFL game.

We suspect the reason for this tight zero (along with the lack of results in the AQI regressions) have something to do with the lack of substitutable alternatives to attending one of eight annual regular season games in each market, and that we are much more likely to see a result from baseball games. While it is unclear if air quality plays an important part in the enjoyment of outdoor leisure, most people would agree that weather does, so these tight zero results on weather may support the theory that NFL games are too unique to reveal the true cost of bad air quality or weather on outdoor leisure.

Table 2.8: Weather Regression Results

	Dependent variable:		
	Log Price		
	Low Climate Temp	Mid Climate Temp	High climate Temp
	(1)	(2)	(3)
1 Day Before	−0.122***	−0.139***	−0.249***
	(0.020)	(0.024)	(0.044)
Temp Diff x 1 Day	0.002	0.001	−0.0001
	(0.001)	(0.002)	(0.007)
Outdoor x 1 day	−0.086***	−0.065**	0.133**
	(0.027)	(0.030)	(0.058)
Temp Diff x Outdoor x 1 Day	−0.0003	0.003	−0.008
	(0.003)	(0.003)	(0.008)
Game FE	Yes	Yes	Yes
Stadium Zone FE	Yes	Yes	Yes
Clusters (Games)	73	104	29
Observations	2,577,677	4,215,765	1,276,324
R^2	0.849	0.823	0.854
Adjusted R^2	0.849	0.823	0.854
Residual Std. Error	0.315 (df = 2577458)	0.288 (df = 4215435)	0.283 (df = 1276188)
Note:			*p<0.1; **p<0.05; ***p<0.01

2.8 Conclusions

This study estimates the impact of poor air quality on willingness-to-pay to attend National Football League games. We use exogenous variation provided by the Northern California wildfires, which led to extremely poor air quality during two NFL games – the November 11th home game for the Oakland Raiders and the November 12th home game for the San Francisco 49ers. We find no effects of AQI increases on listed ticket prices, or on the number of tickets listed, indicating that consumers do not appear to respond to air quality adjustments in the context of outdoor NFL games. We also rule out potential alternative mechanisms that might explain this null result. First, we feel confident ruling out a "lack of transparency" effect, since the Camp Fire in Northern California caused the AQI in the Bay Area to increase by nearly three standard deviations from the mean value found in our sample just before and during the two afformentioned games. Additionally, we also rule out "lack of attention" on the secondary market to time-variant factors that may affect ticket prices, since we see sellers change prices often, and large and significant price responses to the injury announcement of the 49ers star quarterback. We also find a similar null price response to temperature fluctuations in the context of NFL games.

The overarching goal of this study is to estimate a demand response for recreational activities to changes in environmental quality. Because attending NFL games is not a typical recreational activity (like going for a run, hike, or to the park), the external validity of these estimates to other recreation needs to be examined further. Future work will aim to assess these impacts using data from Major League Baseball games, of which there are 81 home games per season (and thus lower opportunity costs of not attending than football, since there are only 8 guaranteed home games for NFL teams). Additionally, we hope to take further advantage of the rich temporal and spatial variation in both of these settings, as these matchups are played across multiple seasons, different locations (including stadiums that are domed versus outside), and reflect different matchup-specific characteristics (like quality of the two teams).

Bibliography

Archsmith, James, Anthony Heyes, and Soodeh Saberian. 2018. "Air quality and error quantity: Pollution and performance in a high-skilled, quality-focused occupation." *Journal of the Association of Environmental and Resource Economists* 5 (4): 827–863.

Bishop, Kelly C, Jonathan D Ketcham, and Nicolai V Kuminoff. 2018. *Hazed and confused: the effect of air pollution on dementia.* Technical report. National Bureau of Economic Research.

Chan, Nathan W, and Casey J Wichman. 2018. *The Effects of Climate on Leisure Demand.* Working paper RFF WP 17-20-REV. Washington, DC: Resources for the Future.

Currie, Janet, and Reed Walker. 2011. "Traffic congestion and infant health: Evidence from E-ZPass." *American Economic Journal: Applied Economics* 3 (1): 65–90.

Egan, Kevin J, Joseph A Herriges, Catherine L Kling, and John A Downing. 2009. "Valuing water quality as a function of water quality measures." *American Journal of Agricultural Economics* 91 (1): 106–123.

Graff Zivin, Joshua, and Matthew Neidell. 2012. "The impact of pollution on worker productivity." *American Economic Review* 102 (7): 3652–73.

———. 2014. "Temperature and the allocation of time: Implications for climate change." *Journal of Labor Economics* 32 (1): 1–26.

Hanemann, W Michael. 1984. "Welfare evaluations in contingent valuation experiments with discrete responses." *American journal of agricultural economics* 66 (3): 332–341.

Kaplan, Scott. 2019. "The Economic Impact of NBA Superstars: Evidence from Missed Games using Ticket Microdata from a Secondary Marketplace." *Working Paper.*

Neidell, Matthew. 2009. "Information, avoidance behavior, and health the effect of ozone on asthma hospitalizations." *Journal of Human resources* 44 (2): 450–478.

Sweeting, Andrew. 2012. "Dynamic pricing behavior in perishable goods markets: Evidence from secondary markets for major league baseball tickets." *Journal of Political Economy* 120 (6): 1133–1172.

Chapter 3

The Effects of an Efficiency Mandate on the Market for Light Bulbs in the United States

3.1 Introduction

That seemingly profitable energy efficiency investments have slowly diffused through society, even though they offer unrealized 'free money', is a long standing problem in the intersection of engineering and environmental economics (Jaffe and Stavins 1994). In fact, in 2009, the famous McKinsey curve, a cost curve estimating energy efficiency investments, labeled residential buildings and lighting the fourth cheapest investment in carbon abatement, with a price of *negative* \$90 an ton of CO_2(Granade et al. 2009).

Expanding energy efficiency adoption is a popular political option because of its promise to help fight climate change while saving households money, but actually finding effective policy prescriptions have been difficult(Gerarden, Newell, and Stavins 2015). Despite this, plans to shrink the energy efficiency gap remain critical components of government plans to fight climate change, such as the (now imperiled) Clean Power Plan(EPA 2005). In fact, widespread demand-side management has been practiced by Californian utilities since the 1970s, even as its payoff has been questioned (Gellings and Chamberlin 1988). In 2010, California utilities spent almost \$1 billion on customer-funded energy efficiency programs (Barbose et al. 2013). After compact florescent bulbs (CFLs), which are nearly 80% more efficient than incandescents, became commercially available in the mid 2000s, utilities and electricity regulators set their sites on making residential lighting the next frontier of demand-side management. In 2007, PG&E spent \$116 million subsidizing CFLs in pursuit of achieving demand reduction targets set by the California Public Utilities Commission (CPUC) (Smith 2008). But actually subsidizing the entire switchover from old, inefficient bulbs to new ones would have be extremely expensive, so governments turned to efficiency mandates.

In this paper, I will assess the effect these mandates had on the residential light bulb market by answering three questions. First, why did no 'efficient' incandescent exist before the efficiency mandate? Specifically, were classic behavioral failures on the part of consumers or market power by producers, or perhaps, even a lack of market failure from the beginning to blame? Second, did the mandate encourage consumers to increase purchases of CFLs? Third, was the policy successful and welfare improving? Section 2 will explain the mandate and the theories on how the market would be affected. Section 3 introduces the panel data of stores from one of the nation's leading retailers. Section 5 introduces the differences models that are used in the analysis. Section 6 presents the results and discusses their meaning, and section 7 will summarizes my findings.

3.2 The Light Bulb Market and the Theory Behind Efficiency Mandates

There are many types of light bulbs, but the classic household upside down pear shaped light bulb is known as an A-shape bulb (as opposed to a flood, reflector, candelabra, or vanity bulb). While CFLs are a different shape altogether, they were designed to replace A-shape bulbs, so for convenience, I will refer to them as A-shaped as well. Before the efficiency mandates, the only types of A-shaped bulbs were CFLs, regular incandescents (of the 40/60/75/100 watt variety), and the slightly more expensive high spectrum incandescents, which are often referred to by their brand names (the biggest in my sample being the Reveal brand). CFLs were almost ten times more expensive than the incandescents, but still rated as a good investment due to their longer life and efficiency. Customers had only two efficiency level choices, and they were far apart on both price and watts.

In 2004, the California Energy Commission (CEC), noting that a 5% efficiency improvement in all incandescents sold would save the same amount of energy as the effect of the millions of dollars the commission and its utilities had spent the prior year supporting CFLs, announced that starting in 2008 they would mandate just that, reasoning that "it would be very cost effective" (CEC 2004). Subsequently, Congress passed the Energy Independence and Security Act (EISA) of 2007, which set more ambitious mandates of approximately 30% for the four classes of incandescents starting in 2012, and allowed California to institute the mandate one year early, which they elected to do (EPA 2011). To meet the first California mandate, California compliant incandescents were introduced. Later, halogen incandescents were introduced which exactly met the new EISA requirements.[1]

What might explain why more efficient incandescents did not exist and that demand for money saving CFLs remained relatively modest? Before we consider the classic explanations for the energy efficiency gap, we should poor some cold water on the theory that adopting energy efficient appliances is always welfare improving. Researchers have routinely found

[1]The mandates prescribed minimum efficiency levels for different levels of brightness, making them attribute-based regulation. Such regulation is likely to lead to bunching around the minimum requirements, and that is exactly what happened (Ito and Sallee 2014).

Table 3.1: Light Bulb Efficiency Mandates

	watts equivalence	2008	2009	2010	2011	2012	2013	2014
California	100	95*	...	...	72	...	...	...
	75	71*	...	...	...	53	...	...
	60	57*	...	...	...	...	43	...
	40	38*	...	...	...	...	29	...
Rest of USA	100	...	...	...	...	72	...	...
	75	...	...	...	...	...	53	...
	60	...	...	...	...	...	...	43
	40	...	...	...	...	...	...	29

*High spectrum bulbs (i.e. Reveal) exempt from 2008 mandate

that the engineering estimates of efficiency investments do not match the realized gains (Allcott and Greenstone 2012). Recent experimental evidence from government financed residential retrofits found energy savings to be much lower than costs, and could not be simply explained away by a rebound effect (Fowlie, Greenstone, and Wolfram 2015). In addition, while engineers often treated CFLs as perfect substitutes for incandescents, it quicly became clear that they weren't. Consumers complained that the light produced by CFLs was of a lower quality, that they took too long to warm up, that they looked strange, that they did not last as long as advertised, and that they needed to be disposed of at a local hazerdous waste site due to containing small levels of mercury. Furthermore, their utility to some consumers likely wanned as they became enmeshed in identity politics. In a lab experiment, Conservatives were less likely to buy an energy-efficient bulb if it was labeled with an environmental message (Gromet, Kunreuther, and Larrick 2013). CFLs briefly became the poster child of liberal environmental policy overreach to the point that conservatives in Congress passed a law prohibiting the EPA from enforcing the mandate (although not repealing it).

It is possible, then, that for at least some consumers, no market failure existed, and that inefficient incandescents were the efficient choice. Testing this possibility will be difficult, but I would expect that large scale stockpiling of incandescent bulbs before mandates take place, an increase in the purchase of bulbs exempt from the mandate after the policy, and a general reduction in demand for bulbs to be signs that some consumers are harmed by a mandate.

Assuming that the engineering estimates are somewhat reasonable for light bulbs, however, leaves us with a classical efficiency gap. Explanations of why the energy efficiency gap persists often blame the behavioral biases of consumers, namely bounded rationality, reference-dependent preferences, biased beliefs, and inattention (Gillingham, Newell, and Palmer 2009) (Gillingham and Palmer 2014). If consumers are not willing to adopt a new, promising technology due to prospect theory, this is likely to lead to loss aversion, anchoring, and status quo bias (Shogren and Taylor 2008). In other words, if consumers are used to buying 60 watt bulbs, they will be unlikely to try CFLs or a newer more efficient incandescent, even if they would likely be better off. This could explain why many proponents predicted that efficiency mandates that eliminated consumer's default choice would encourage many consumers to convert to CFLs en mass(Smith 2008). To test this hypothesis, I will examine if there was a significant increase in deamnd for CFLs after the mandate.

These classic behavioral mistakes could also explain why a market for more efficient incandescents never materialized before the mandates, but a more 'rational' explanation is possible as well. If making efficiency calculations is sufficiently complex and difficult, it may be rational for consumers to ignore them (Sallee 2014). Light bulbs may not seem like a good candidate for rational inattention, since common household incandescents have very few attributes other than their brightness (measured in lumens), efficiency (measured in watts), color, and shape. In fact, consumers mostly identify bulbs by their efficiency. But that may have been the problem: consumers mistake the efficiency of the light bulb as a measure of its brightness. This is so ingrained that more efficient bulbs are not labeled with their lumens, but with the efficiency of the old bulbs that have an equivalent brightness. A halogen bulb wouldn't label itself as a 43 watt bulb with 800 lumens, but a 43 watt bulb with the equivalent brightness of a 60 watt bulb. It is easy to see how this could get confusing for anyone picking between a \$0.75 bulb and a \$1.25 halogen and simply wants to make sure they are not sitting in a dim room. Other behavioral biases may effect purchases of light bulbs.

If bounded rationality or rational inattention are to blame for the lack of more efficient options, we would expect little stockpiling and no drop off in overall demand for bulbs, as the mandate corrects for mistakes people are making, or people are made better off without having to think about it. Seeing demand for exempt bulbs increase, however, does not necessarily disprove this theory, however, as behavioral errors or inattention could explain loyalty to high watt bulbs as well.

All these theories deal with the consumer, but the mandate could be correcting a market failure on the supply side. Fischer 2005 imagines a market for appliances with limited competition and two types of consumers: type one, who have a higher demand for efficiency, and type two, who have a lower demand for efficiency. Like the classic price discrimination case, the producer captures all of the low type's consumer surplus, but is unable to price the more efficient good high enough to capture all of the high type's surplus because of an incentive compatibility constraint. Fisher's contribution is to notice that the producer can increase the price she charges the high type by offering a less than optimal efficiency to the low type since Fisher shows:

$$\frac{\mathrm{d}p_1}{\mathrm{d}\phi_2} = (\beta_1 - \beta_2)g > 0$$

Where p_1 is the price of the more efficient appliance, ϕ_2 is the energy intensity of the less efficient appliance, β is the cumulative discount factor for the utility flow over the lifetime of the appliance (with $\beta_1 > \beta_2$), and g is the price of electricity. An efficiency mandate on the low good would fix this market failure. This hypothesis could be supported if an efficiency mandate that rises the price and lowers the energy intensity of incandescents produced in a less than fully competitive market does not change the demand for either good but results in a lower price for the CFL.

This story could fit well with the final mandate I will examine that takes place in 2013, when all inefficient incandescents are pulled from the market in California and replaced by more expensive halogen incandescent. In the first two mandates, however, producers respond by offering dimmer bulbs with fewer lumens[2]. This means that the value (v_i in Fischer's notation) of the appliance in these occasions is now a function of the energy intensity, with $v_i'(\phi_j) > 0$. This means that the change in the CFL price is ambiguous since:

$$\frac{\mathrm{d}p_1}{\mathrm{d}\phi_2} = \beta_1\big(g - v_1'(\phi_2)\big) - \beta_2\big(g - v_2'(\phi_2)\big)$$

If the reduction in utility from the value of a dimmer bulb matches the energy savings, then the incentive compatibility constraint does not change, so the price of CFLs should not change.

R&D and learning by doing spillovers are additional supply side explanations for the lack of efficient incandescents. This is a particularly hard hypothesis to test, but lower prices for halogen incandescents over time would be consistent with this story.

3.3 Store Panel Scanner Data

In order to investigate how the mandates affected the market for light bulbs, weekly store level data was obtained for one of the nation's largest big box retailers via the Nielsen retail scanner dataset. This dataset includes weekly upc level summations of the number of bulbs sold and their average price at every store by some of the nations largest grocery stores, pharmacies, and big box chains. Instead of using all stores, I have limited the scope of the investigation to just the largest store in the Nielsen sample[3]. Since the scanner data set is not a random selection of all stores, it would never be possible to get a representative sample of all light bulb purchases, limiting the advantage of using more retailers. Instead, by only using one retailer, I can ensure parallel trends for my differences in difference models more easily. The retailer, for the most part, only sold General Electric brand light bulbs and their own generic brand[4], making it a good candidate to study market power. Figure 3.1, which

[2]In 2008, in response to the 5% mandate, producers created bulbs that were 5% dimmer. In 2011, when the 95 watt bulb was made illegal, no halogen bulbs were ready, so the low option was the 71 watt, dimmer incandescent

[3]As a robustness check, I conducted the same analyses with another mass marketer that had significant (but lower) sales in both California and the rest of the country. This retailer had approximately 1% of the national and California market, and sold Sylvania bulbs (the nation's second largest brand with a market share of 17%) almost exclusively. I found little evidence of different outcomes, so I am reasonably confident the following results are not limited to a single retailer

[4]During the study period, 92% and 89% of incandescents and CFLs were GE branded bulbs, with almost all of the rest coming from the store's in house brand. In the household panel, GE was also the largest

Table 3.2: Bulbs per Store per Week

			All Incandescents				CFLs			
			Bulbs		Price Bulb ($)		Bulbs		Price Bulb ($)	
Year	State	Stores	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
2007	CA	226	488.3	314.4	0.44	0.05	78.1	62.2	3.84	0.51
	USA	1379	507.4	327.4	0.43	0.05	64.3	46.8	3.87	0.47
2008	CA	243	369.2	258.5	0.45	0.07	73.6	57.9	3.67	0.58
	USA	1458	407.8	265.9	0.43	0.05	67.9	50.2	3.63	0.53
2009	CA	245	347.6	209.3	0.51	0.09	78.4	57.5	3.42	0.67
	USA	1503	393.4	246.4	0.47	0.08	71.4	51	3.43	0.6
2010	CA	249	299.6	190.3	0.63	0.09	75.4	57.6	3.65	0.67
	USA	1502	368.8	246.9	0.57	0.09	73	63.4	3.63	0.72
2011	CA	254	276.7	193.2	0.98	0.25	77.4	55.3	3.98	0.62
	USA	1513	380.4	248.2	0.76	0.14	71.8	53.7	3.97	0.65
2012	CA	258	205.6	126.9	1.29	0.2	69.5	50.6	4.62	0.67
	USA	1525	296.9	193.1	0.91	0.15	63.2	46.7	4.6	0.69
2013	CA	263	175.3	110.2	1.63	0.26	71.8	51	4.55	0.63
	USA	1532	255.7	166.9	1.06	0.18	65.2	47.9	4.53	0.62

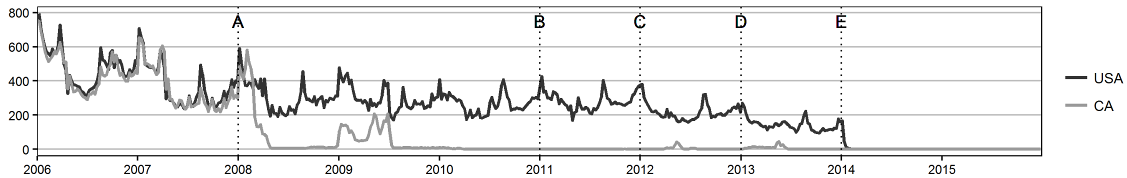
shows the retailer's sales of different bulbs over time, shows that when sales occur, they tend to occur nationally, and that the average store in California, before the policy was enacted, sold very similar numbers of bulbs as the average outside of California[5]. Very few retailers in the Nielsen dataset have such a national footprint, so by adding in more retailers, I would inevitably be comparing very different types of shoppers and prices. As a final benefit, by limiting the sample to one store, the number of UPC codes that must be categorized by the researcher is reduced from thousands to hundreds, ensuring accuracy of the results.

The drawbacks of using a single retail chain are related to external validity. If different types of retailers or consumers react differently to the mandate, that will be missed in my analysis. While the big box store used is by far the largest in the Nielsen scanner data, it is not the largest bulb seller in the country. The household panel during the same period (which includes purchases at all stores) indicates the store used was the forth largest bulb retailer in California and the seventh largest in the country, with a market share of 7.4% and 2.8%, respectively. In addition, by not incorporating household data (which is very difficult to obtain on a large enough scale given the infrequency that households buy light bulbs), we will be unable to conduct individual purchasing decision models.

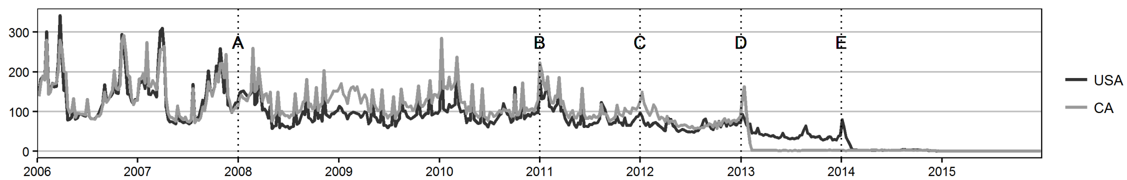
brand, accounting for 45% and 30% of incandescent and CFL bulbs sold, making it reasonable that it could exercise some market power.

[5]This is consistent with the promotional patterns in section 1 of this thesis, as promotions are never set at the store level

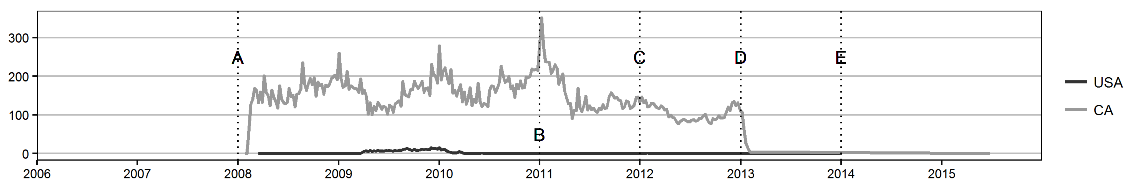
Figure 3.1: A-Shape bulbs sold per week, per store.



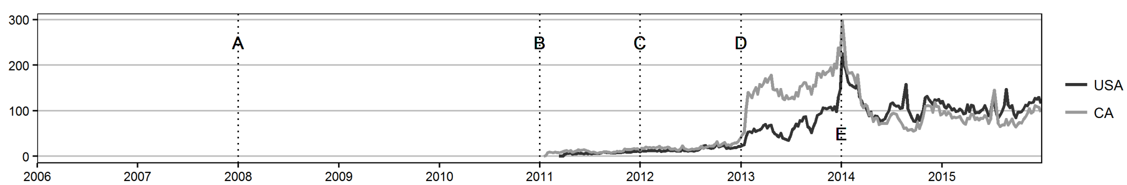
(a) Regular Incandescents



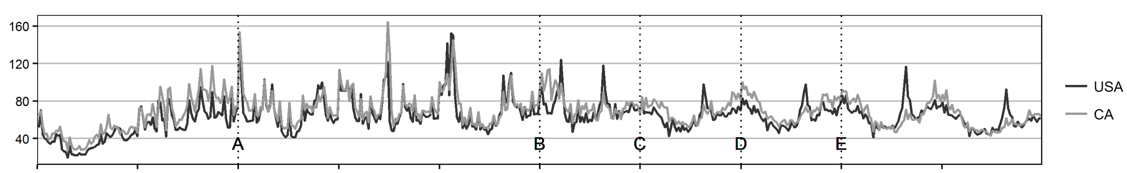
(b) Reveal Brand Incandescents



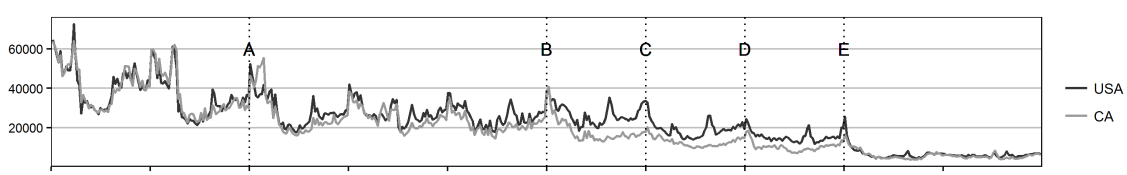
(c) California Compliant Incandescents



(d) Halogen Incandescents



(e) CFLs



(f) Sum of Watts from Incandescents and CFLs Sold per Store

A) (2008) CA: Regular incandescents replaced by California compliant incandescents...B) (2011) CA: 95w -> 72w...C) (2012) CA: 71w -> 53w USA: 100w -> 72w...D) (2013) CA: 57w -> 43w, 38w -> 29w USA: 75w -> 53w...E) (2014) USA: 60w -> 43w, 40w -> 29w.

3.4 Modeling the Changes in the Light Bulb Market

Simple Difference Model

With a panel of observations of our outcomes of interest (both number of bulbs sold and the price charged by the retail chain) and a series of single policy changes, the simplest model to construct would be a two period difference in difference. The margins of variation of interest are the state that store i is in ($s \in$ CA or USA) and the time period of observation ($t \in$ Pre or Policy). The Nielsen data observations are at the week (w) and store (ς) level, so week and store fixed effects can be used, meaning the basic difference in difference regression model is:

$$Y_{w\varsigma} = \alpha + \beta D_{st} + \gamma \mathbb{1}(w_t = 1) + \delta \mathbb{1}(\varsigma_s = 1) + \varepsilon_{w\varsigma}$$

Where Y is our outcome of interest (either number of bulbs sold or price) and D_{st} is the interaction between state and period. As in any difference in difference model, the key identifying assumption is parallel trends: absent the treatment (t), we need to expect our outcome of interest to have the same trajectory over the range of time studied. As mentioned, limiting the study to just sales from a single retailer has drawbacks, but its main advantage is to support the parallel trends assumption. Figure 3.1 shows that, especially before the first policy change in 2008, average stores in California and the rest of the country sold a similar number of incandescent and CFL bulbs. The frequent and brief spikes in quantity correspond to short term price reductions that, happily, are national in nature. More generally, sales increase in the winter, when it is darker, and decrease in the summer in both California and the rest of the country.

I will more formally test the parallel trends assumption with a dynamic differences model, but there are three visible concerns. First, when I analyze the policy changes in 2011 and 2013, California stores no longer sell regular incandescents, instead offering the 5% more efficient California incandescents, which may mean subtle differences in the trends in bulb purchases. Second, beginning in 2007, stores in the rest of the country experience large spikes in sales at the end of August, but there is much less of a corresponding spike in California.[6] To deal with this seasonal variation, I introduce a month by state fixed effect in some robustness checks.[7]. Third, California demand for CFLs is higher in the years before the policy[8]. While the trends seem to still be parallel, since CFLs last longer, early purchasers of CFLs could have lowerer demand in the future while they wait for their bulbs to burn out.

[6] Examining the data, I have not been able to conclusively determine why this spike exists. There is no major state, August variation in price and an indicator variable for featured display and sale prices available in some stores in the panel is no more frequent than during the rest of the year. I hypothesize that this period correlates with the back to school shopping period, but it is not clear why there is a much smaller spike in sales in California.

[7] My difference in difference specifications use two years of data each, so a month (not month x year) x state fixed effect is not collinear with the week fixed effect (which is a week x year specific indicator)

[8] This may possibly have been due to the CA utilities promotion of CFLs, but I have identified the brands that they specifically subsidized and they are never sold at my retailer of interest

Assuming that these potential problems with the parallel trends assumption only cause marginal bias or can be controlled by additional fixed effects such as the month x state variable, and assuming that there are no other state specific factors that vary over the time frame of the difference model, the model should be well identified. Just using a single time variable is probably unrealistic, however. Because the mandates only apply to bulbs produced after January 1st, I observe that stores sometimes change the price of their stock of discontinued bulbs to sell them off in the first month or two of the new year and make room for the new bulbs. Furthermore, there is evidence from the Nielsen household panel that some consumers stockpiled soon to be unavailable bulbs. Examining each policy roll out, I identified a month or two of a 'sale' time period where there was a drastic price change in soon to be or currently illegal bulbs or where there was still significant sales of the stores' stock. This potential stockpiling time period is policy relevant, as well, as it could suppress sales in the following months. Thus, the final difference model used for most specifications is:

$$Y_{w\varsigma} = \alpha + \Sigma_{t=1}^{2}\beta_t D_{st} + \gamma_1 w_t + \gamma_2 \varsigma_s + \varepsilon_{w\varsigma} \tag{3.1}$$

Our observations, weekly sales data from stores in a single retailer across the country, are likely to exhibit correlation within certain groups. For instance, if there is a nationwide sale one week, the week fixed effect will correct for the increased level of sales, but sales across the country are obviously correlated with each other that week. In addition, a national retailer's supply chain is likely to experience different conditions across the country, and economic shocks or consumption tastes are likely to effect demand in geographically correlated ways. Ignoring these correlations will result in standard errors that are too small. To correct for this, I use cluster standard errors at the week and market level. Markets roughly correspond to metro areas, and are defined by Nielsen as the 210 television markets across the country (of which, 191 have at least one store in the sample). While in many cases, state is a higher level group than metro, that is not always the case, nor is state obviously a better cluster. For instance, the economy, consumer tastes, and supply chain in New York City are more likely to be correlated with Newark than with Buffalo. Indeed, state clustered errors are smaller than metro clustered errors, so metro clustered errors are likely to be the more conservative choice.

In general, clusters should be applied at the level of treatment or even one level above, but that is impossible in this case, since the treatment is California and not California. If California was like east coast states whose boarders bisect metro areas, I could examine stores on each side of the boarder and cluster at the metro area. Since this option is not available to me, I am still concerned that the standard errors are still too small. In order to examine this possibility, I follow Bertrand 2004 and use the untreated states as placebos. For most difference regressions (see table 3.3, 3.5, and 3.7), I drop all treated (California) observations, and then rerun the model using each of the 11 states who have more than 50 stores as placebos. The tables report the frequency that the state x policy variable of interest is significant at the 99% level. Unfortunately, I find a significant coefficient estimate on the variable of interest about 10% of the time, which suggests that standard errors are still too small. Thus, results that are not significant at at least that level should be viewed with skepticism.

A final modeling consideration is using characteristics of the light bulbs themselves as fixed effects. The data are reported not at the store/week/light bulb type level but at the store/week/upc code level. UPC codes, unfortunately, cannot be used as fixed effects, since each time a policy is implemented in California, all the UPC codes change. It is possible, however, to control for the watts equivalence.[9] As a robustness check, I have run all models presented with watts equivalence fixed effects. All models that predict bulbs sold have nearly identical coefficient estimates, suggesting this does not add any new information to the model, other than multiplying the data by four (as there are now four observations for each store and week). Models that predict price do have different results, but only in cases where specific watts classes of bulbs are regulated out of the market, and these changes are almost certainly the result of bias from the few stores that ignore the ban or misclassified bulbs.[10] Thus, using these fixed effects adds nothing to the model, but may artificially reduce standard errors or introduce bias. The preferred method of seeing the effects of the mandate on different types of bulbs is to rerun the model, each time restricting it to only certain classes of bulbs.

Triple Difference Model

Time variant state-specific shocks could still be a real problem, however. 2008 in particular was a time of great economic disruptions tied to the housing market, and were thus strongly state dependent. In order to net out much of those changes, we can introduce a third difference by using sales of non A-shaped bulbs (which were largely candelabra shaped decorative bulbs, flood lamps, globe shaped vanity bulbs, outdoor utility bulbs, etc...Christmas string lights and night lights were not used due to their extreme seasonal variation). Since the mandates largely left these types of bulbs unregulated, and since they are poor substitutes for A-shaped bulbs.[11] Using a to denote A-shaped bulbs from others, our model for the triple difference is as follows:

$$Y_{w\varsigma a} = \alpha + \Sigma_{t=1}^{2}\beta_t D_{sta} + \gamma_1 w_t + \gamma_2 \varsigma_s + \gamma_3 a + \gamma_4 \varsigma_s \text{ x } a + \Sigma_{t=1}^{2}(\delta_t(t \text{ x } a)) + \Sigma_{t=1}^{2}(\zeta_t(t \text{ x } a \text{ x } \varsigma_s)) + \varepsilon_{w\varsigma a} \tag{3.2}$$

[9] A 60w regular or Reveal bulb, a 57w California compliant bulb, a 43w halogen bulb, and a 13w CFL all have a watts equivalence of 60

[10] To see this, consider the 2011 mandate against 100w bulbs in California and a model estimating the price of all incandescents. After the ban, most stores will report no 100w sales, but a few may report a very small number of sales at a very high markup. Since the price at all of the compliant stores is missing, the coefficient on the 100w fixed effect is now very large. Interacting the fixed effect with the policy indicator would defeat the purpose of the model: estimating price changes for all incandescents

[11] the degree of substitutability is somewhat arguable. Candelabra and most vanity bulbs have a different base diameter, but flood lights (the largest class of 'other' bulb) can fit in the same sockets. Since they are considerably more expensive, it is unlikely that a consumer would use a flood light rather than a regular A-shaped bulb in a lamp, for instance, but customers may use the cheaper regular bulbs in a place where they usually use flood lights in response to an economic shock

As a robustness check, all models presented were also conducted with triple differences. Changes, except when mentioned, were minor, which suggests that time variant state dependent shocks are not a source of major bias. Because the triple differences largely confirm the simpler models, in order to ease the interpretation of the size of the effects of the mandate on the quantity and price changes, most models are presented in the difference in difference format.

Dynamic Difference Model

The mandates apply to bulbs *manufactured* after a certain date, so the effect at the retail level may lag behind by a month or two. In addition, after a new product is introduced, a supply chain may take a few months to respond to demand and price and quantity may fluctuate somewhat. Because I have no instrument for price, it is reasonable to be concerned how price is effecting quantity and vice versa. By treating all observations the same in either the entire year before or the entire year after the policy change, the difference in difference model makes it hard to see how price and quantity may be moving together. In order to better see these dynamic effects, I use a version of equation 1 that instead of having two time periods (pre and policy) or three (pre, sale, and policy), has a month dummy interacted with the variables of interest, so the dynamic effect can be seen over time.

$$Y_{w\varsigma} = \alpha + \Sigma_{t=1}^{24}\beta_t D_{st} + \gamma_1 w_t + \gamma_2 \varsigma_s + \varepsilon_{w\varsigma} \tag{3.3}$$

Equation 3 can also be expressed using the triple difference model, but due to a lack of major differences, I present only the difference in difference version for ease of interpretation. Since there are only four or five weeks per month, and month/California is now the treatment level, I only cluster standard errors by market. In addition, including a state x month fixed effect is now impossible, so the seasonal differences between California and the rest of the country will be pronounced in the output (this is most evident every August), so care should be taken in looking at year over year changes when interpreting figures 3.2, 3.3, and 3.4.

Price and Quantity Simultaneity

Without an instrument for price, it is highly problematic to include price in a model predicting quantity. As a robustness check, I have included one model that does that for each mandate, but I am highly skeptical of those models. Inspecting models that separately predict quantity and price should be able to still shed plenty of light on the effect of the mandates. If quantity demanded goes down and price goes up after a mandate roll out, it may be impossible to tell if people bought less light bulbs because they have a lower value for the new bulb or because the price went up, but the first order question to a policy maker is what happened because of a policy, and not necessarily the exact mechanism. Without an instrument, I cannot make strong statements about the dollar value customers place on efficiency, but I can say more generally how a mandate effects the overall market.

3.5 Results and Discussion

In this section, I will examine each of the three mandates of interest. First, in January 2008, California instituted its 5% mandate, switching from regular incandescents to California incandescents, but exempting slightly more expensive full spectrum A-shape incandescents (Reveal brand). Second, in January 2011, California instituted the first leg of the EISA by requiring all 100 watt equivalent incandescents (California and Reveal bulbs) consume only 72 watts, but the industry had yet to begin widespread production on the halogen replacements, so no 100 watt equivalent incandescents replaced them at the time. Third, in January 2013, the full EISA mandate came into effect in California, making all incandescents that did not meet the mandate (basically, all non-halogen incandescents) illegal in the state. This was also the first time that halogen incandescents of all four watt equivalents were widely available in California stores (see figure 3.1). Meanwhile, in the control group, the rest of the United States, 40 and 60 watt bulbs were still available, though 75 watt bulbs were made unavailable that year. For brevity, I will not examine the 2012 roll out when 75 watt incandescents were removed from the market in California and 100 watt bulbs were removed in the rest of the country. In all three cases, I examine difference in difference results as well as dynamic difference models.

2008, 5% CA Mandate

The first mandate in 2008 is of interest because it is the only mandate where the markets in California and the rest of the country were identical previous to the mandate, and it is the only case of California acting completely independently of the rest of the country. In addition, of all the mandates, this had the largest available loophole to consumers, who could simply pay a little bit more to buy a Reveal bulb. In December 2007, an average regular incandescent cost \$0.37 in California while the Reveal bulb cost \$0.70, but a CFL cost \$3.92 (see table 3.4). In order to comply with the mandate, manufacturers simply made each class of light bulb about 5% dimmer, so no new technology was introduced, and the appliances were made slightly (although probably imperceptibly) worse.

Table 3.3 reports the results for the difference in difference models of most interest. California stores were slow to comply, selling regular bulbs at a normal price until mid January, and then selling a lot of regular bulbs at a steep discount for another month until California compliant incandescents replaced them at the end of February. The sell-off resulted in an increase of incandescent sales of around 200 bulbs per store per week. Because this was associated with a large price decrease (panel h of figure 3.2), it is likely that this spike in sales was the retailer clearing their inventory and not primarily caused by consumers anticipating the mandate (which had almost no coverage in the media) and stockpiling to avoid it.

What seems to be clear from the models in table 3.3 is that after the policy was implemented beginning in March, Californian stores sold a bit more than 40 fewer bulbs a week (a decrease of about 10%). The triple difference confirms that this was not due to a time variant shift in Californians' demand for light bulbs. Not only that, CFL sales also go

Table 3.3: 2008 California 5% Efficiency Mandate

	Bulbs per Week per Store							
	All A-Shaped Bulbs			Triple Diff	CFLs	Incand	Reveal	Other Incnd
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sale x CA	198.59***	161.04***	166.94***	27.06***	1.71	196.90***	−1.40	198.32***
	(28.41)	(25.23)	(25.98)	(5.07)	(2.66)	(30.32)	(5.24)	(27.96)
Policy x CA	−43.31***	−42.36***	−36.11***	−2.48	−8.90***	−34.39***	33.75***	−68.12***
	(12.04)	(9.74)	(11.07)	(6.19)	(2.54)	(12.02)	(3.69)	(10.41)
Price			−163.06***					
			(17.00)					
A-Shape Bulbs				189.60***				
				(18.43)				
CA x A				−3.72				
				(14.27)				
Sale x A				3.45				
				(20.92)				
Policy x A				−89.47***				
				(17.77)				
Sale x CA x A				171.39***				
				(23.81)				
Policy x CA x A				−41.13***				
				(5.89)				
Month x St FE	*No*	*Yes*	*No*	*No*	*No*	*No*	*No*	*No*
Week & Store FE	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
St Placebo Sig at 99%	1/11	–	2/11	3/11	1/11	1/11	2/11	1/11
Observations	161,376	161,376	161,376	322,752	161,356	161,376	161,308	161,375
R^2	0.81	0.82	0.82	0.78	0.73	0.79	0.69	0.75

Note: *p<0.1; **p<0.05; ***p<0.01
Standard Errors are clustered at the market and week level.
Pre period is wk ending 01-01-07 to 01-19-08. Sale is 01-26-08 to 02-23-08. Policy is 03-01-08 to 12-27-08

down as a result of the policy by about 10% a week, and this CFL result is also robust to a (not presented due to space considerations) triple difference model. The components of the reduction in incandescent[12] sales suggests that each store sold almost 70 fewer non Reveal, incandescent bulbs than they would have sold if they were able to continue selling regular bulbs, while California stores sold more than 30 *more* Reveal bulbs, suggesting that at least some customers avoided the efficiency mandate by buying more expensive, less efficient bulbs.

Table 3.4: December 2007 Week Store Avg

	California		Rest of USA	
	Bulbs	Price ($)	Bulbs	Price ($)
All Incandescents	424.5	0.45	496.4	0.44
All CFLs	78.3	3.92	66.9	3.93
Reveal Incandescents	109.2	0.7	113.9	0.7
Non Reveal Incandescents	315.2	0.37	382.5	0.37

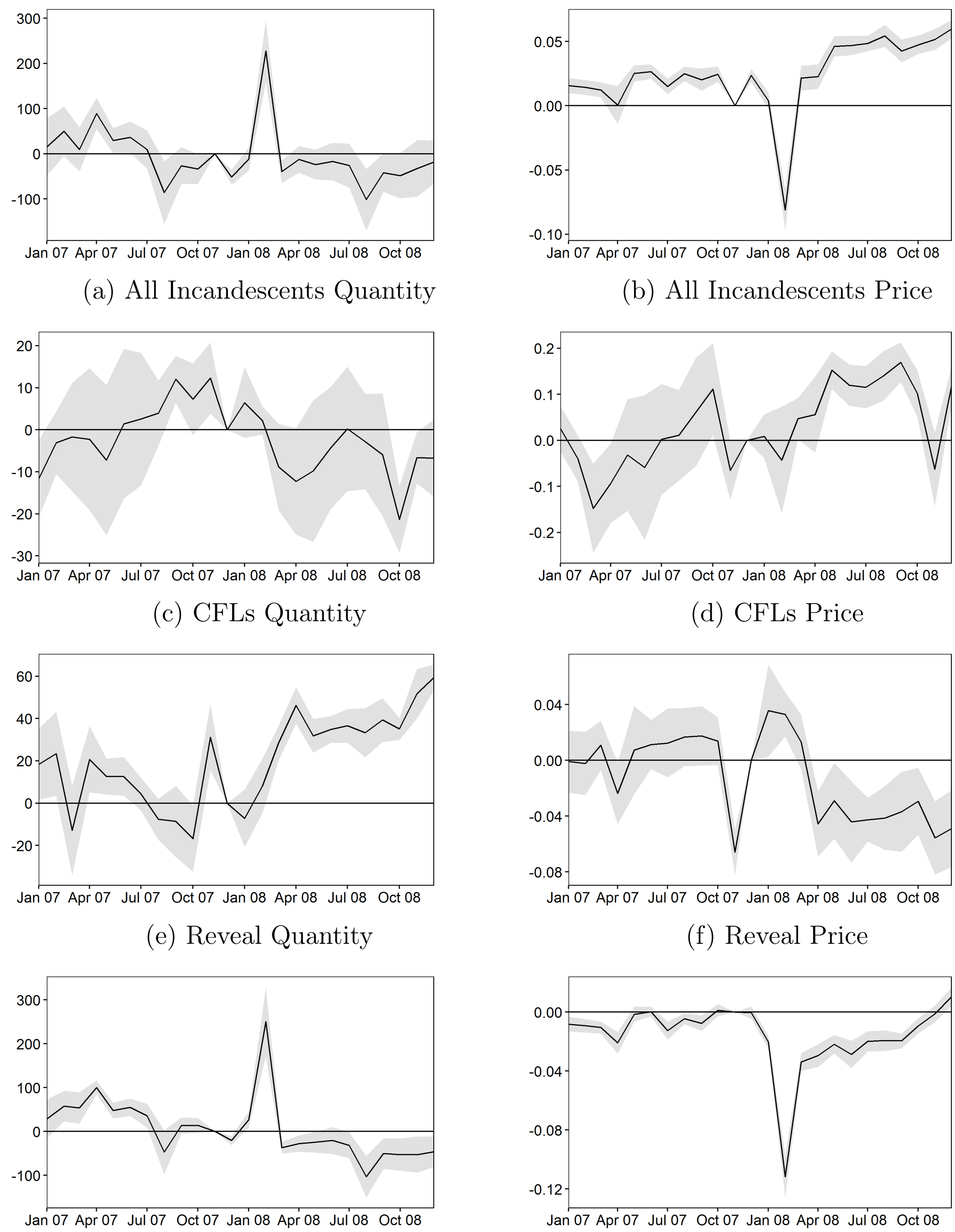
Turning attention to the dynamic difference models presented in figure 3.2, there is little proof of market power being corrected for. Even though demand for the low end bulbs goes down (panel g), the price is lowered in comparison to the regular bulbs being sold elsewhere for most of the rest of the year (panel h). Two possible explanations for this price decrease include cheaper production costs for slightly dimmer bulbs, or, because the bulbs made people worse off, price needed to be lowered to keep consumers in the market. The price of Reveal bulbs also went down after the mandate (panel d), again by a significant but small amount[13]. This raises the possibility that some or all of the increase in Reveal sales was do to a lower price, and not consumer preference for the less efficient bulb. When a difference in difference model is run to predict Reveal bulbs sales controlling for price, the increase in Reveal sales remains, suggesting that this increase may be independent of the price decrease, but there are serious simultaneity issues with such a model. Still, a 30% increase in demand due to a price reduction of only 6% would imply a suspiciously high elasticity. It is thus likely that at least some consumers were paying more to avoid the efficiency mandate.

The drop in CFL sales is also a bit mysterious, although the months with the largest decreases correspond to months with significant price increases. If the California mandate was correcting a market power fed underprovision in inefficiency, we would expect the price of CFLs to go down, not up.

[12]as a reminder, other incandescents(model 8) includes both the regular 40/60/75/100 incandescents and the new California compliant bulbs. The Reveal bulbs in model 7 and the other incandescent bulbs in model 8 make up all incandescents (model 6). Adding together CFLs (model 5) and all incandescents together results in all A-shaped bulbs (models 1-4)

[13]it is unclear why this happened

Figure 3.2: Differences in Store per Week Sales, California, 2007-2008



(a) All Incandescents Quantity

(b) All Incandescents Price

(c) CFLs Quantity

(d) CFLs Price

(e) Reveal Quantity

(f) Reveal Price

(g) Non Reveal Incand Quantity

(h) Non Reveal Incand Price

Note: Shaded areas are the 99% confidence intervals using standard errors clustered at the metro level. Results are from equation 3 using the category of bulb and either number of bulbs sold or average price on the left hand side. The omitted month for CFL and Reveal bulbs is Dec 2007 while the omitted month for all incandescents and non-Reveal incandescents is Nov 2007 (to avoid any stockpiling or sales). All incandescents are regular, CA compliant, and Reveal bulbs.

There are a few reasons why the drop in overall consumption may not be attributable to the policy. First, the large sell-off of the regular bulbs could have suppressed demand later in the year. Second, the parallel trends assumption may not hold. Panel g shows that demand for regularly priced incandescents was already falling and that the downward trajectory may have been a simple continuation of that trend. Differences in the demand for Reveal bulbs were far more evenly distributed around zero before the policy, so the increase in that demand is more robust. Finally, while this retailer is a very large supplier of light bulbs, it by no means had a monopoly, and there exists some evidence in the Nielsen data that other chains were slower to make the change to California complaint bulbs, suggesting some consumers may have shopped elsewhere for their now illegal bulbs.

This evidence allows us to conclude that this mandate was not correcting for an under-provision of efficiency due to market power. The switch to more expensive bulbs that were omitted from the mandate could be evidence that consumers were harmed by the mandate, and that behavioral failures or rational inattention were not to blame for buying inefficient bulbs, but it could also be evidence that these biases are strong enough that about 10% of consumers will pay nearly double to avoid having to make a profitable change. Many consumers default decision was changed, but this jolt did not do anything to increase CFL sales, which refuses the optimistic hypothesis that anchoring or status quo bias were significant forces in holding back CFL sales.

2011, CA EISA 100w Mandate

The conditions present during the 2011 roll out of the efficiency mandate for 100 watt equivalent bulbs in California presented a perfect opportunity to study the optimist theory that anchoring was preventing widespread adoption of the high efficiency CFLs. Because the halogen replacements were not yet widely available on the market for all of 2011 (and 2012), the only 100 watt equivalent bulbs for sale in California were CFLs.

Despite these conditions, I find no effect of the mandate on demand for 100 watt CFLs. The simple difference in difference (table 3.5, model 5) finds a very tight zero effect, and panels c and d show strong parallel trends of both quantity and price before the policy. After the policy, there is a small increase in demand (on the order of 5 bulbs a week compared to the over 80 100w incandescents a week that had been sold in December 2010) followed by a decrease and then nearly no difference in demand is detected for the rest of the year.

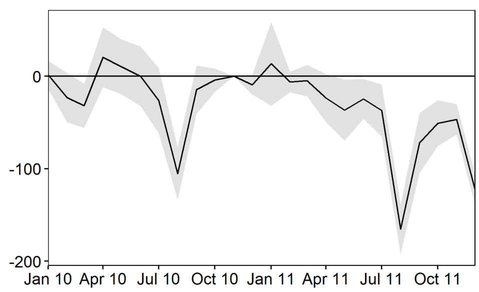
From the other difference models in table 3.5, and from panel e and g of figure 3.3, it is easy to see what happened instead of a drastic increase in the demand for CFLs. After the mandate, the quantity sold of 100 watt incandescent bulbs was reduced by approximately 100%, while there was an accompanying drastic increase in the number of 75 watt bulbs sold. In the simple difference in differences, in table 3.5, I find that the increase in 75 watt bulbs is not big enough to replace all of the missing 100 watt sales, but carefully studying the results from the dynamic differences tells a different story. First, in both panel e and g, 100 watt equivalent incandescents and all other incandescents display strong parallel trends

Table 3.5: 2011 California 100 Watt Efficiency Mandate

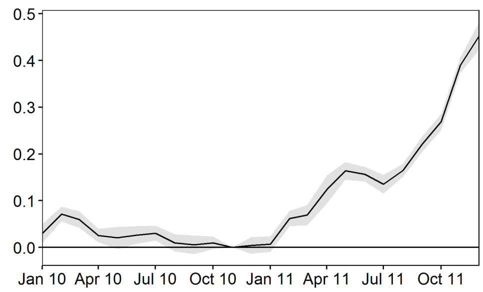
	Bulbs per Week per Store							
	All A-Shaped Bulbs			Triple Diff	100w CFLs	100w Incnd	75w All	40/60w All
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sale x CA	55.71**	47.22***	54.82**	53.21***	5.83***	−26.97	42.11***	45.72***
	(23.32)	(14.19)	(22.06)	(8.58)	(0.99)	(16.43)	(10.04)	(14.59)
Policy x CA	−34.75***	−33.84***	−2.93	−14.09***	0.05	−74.46***	32.16***	10.43
	(9.09)	(6.24)	(9.14)	(4.63)	(0.40)	(4.58)	(3.01)	(7.13)
Price			−151.63***					
			(14.24)					
A-Shape Bulbs				161.61***				
				(12.97)				
CA x A				−72.72***				
				(11.96)				
Sale x A				76.91***				
				(15.19)				
Policy x A				8.28				
				(11.38)				
Sale x CA x A				2.35				
				(14.35)				
Policy x CA x A				−20.92***				
				(7.82)				
Month x St FE	*No*	*Yes*	*No*	*No*	*No*	*No*	*No*	*No*
Week & Store FE	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
St Placebo Sig at 99%	1/11	–	0/11	0/11	1/11	1/11	1/11	3/11
Observations	175,707	175,707	175,707	351,414	169,294	179,068	175,523	175,707
R^2	0.79	0.81	0.80	0.77	0.60	0.67	0.72	0.73

Note: *p<0.1; **p<0.05; ***p<0.01
Standard Errors are clustered at the market and week level.
Pre period is wk ending 01-03-10 to 01-01-11. Sale is 01-08-11 to 02-05-11. Policy is 02-12-11 to 12-03-11

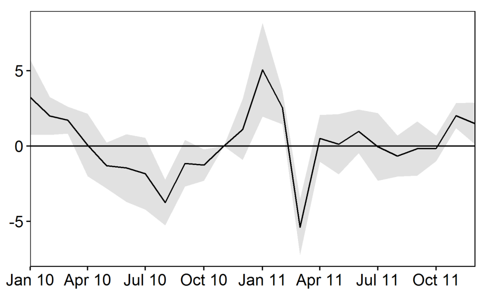
Figure 3.3: Differences in Store per Week Sales, California, 2010-2011



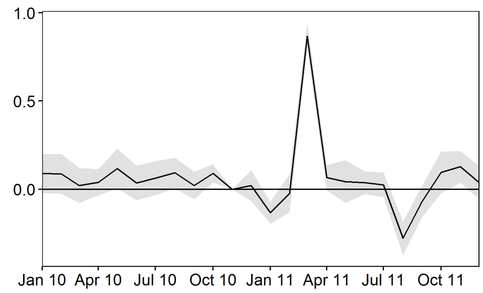
(a) All Incandescents Quantity



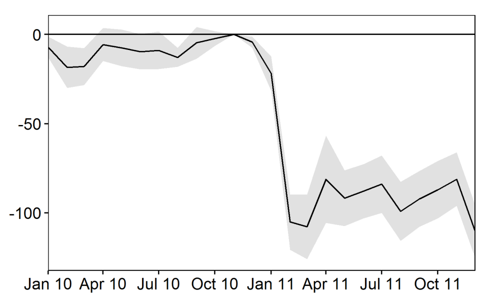
(b) All Incandescents Price



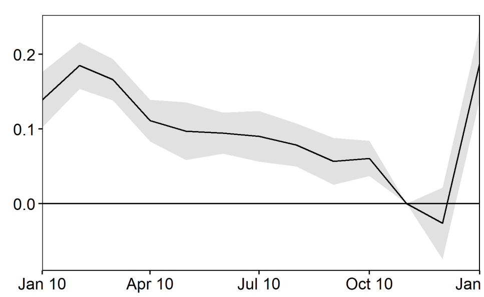
(c) 100w CFLs Quantity



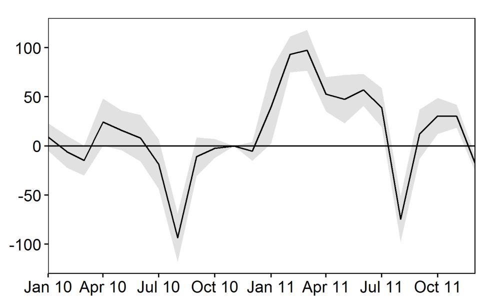
(d) 100w CFLs Price



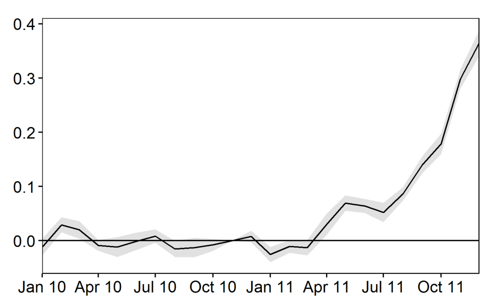
(e) 100w Incandescents Quantity



(f) 100w Incandescents Price



(g) 40/60/75 Incand Quantity



(h) 40/60/75 Incand Price

Note: Shaded areas are the 99% confidence intervals using standard errors clustered at the metro level. Results are from equation 3 using the category of bulb and either number of bulbs sold or average price on the left hand side. The omitted month is Nov 2010 to avoid any stockpiling or sales. All incandescents are regular, CA compliant, Reveal, and halogen bulbs.

Table 3.6: December 2010 Week Store Avg

	California		Rest of USA	
	Bulbs	Price ($)	Bulbs	Price ($)
All Incandescents	330.6	0.69	396.5	0.65
100w CFLs	21.8	4.2	17.3	4.25
100w Incandescents	88.7	0.56	74.7	0.67
40/60/75w Incandescents	241.9	0.73	321.8	0.64

in sales before the policy[14] and then in February, the decrease in sales in 100w incandescents is almost perfectly matched by an increase in sales of lower watt bulbs. Clearly, almost all customers who would have bought a 100w incandescent settled for a 75 watt bulb (panel a of figure 3.3 shows how little effect the policy has in February). Since there is such a strong behavioral response here, the tight zero response in the CFL market by contrast is all the more definitive.

After February, however, the quantity of all incandescents sold declines, but this happens in tandem with a steady and drastic 50% price increase of incandescents sold in California. Furthermore, there is a good amount of circumstantial evidence that this price increase was not related to the mandate. First, the price increase on non 100 watt equivalent incandescents did not start until April. Second, while the retailer (or manufacturer) could be raising the prices of the 75 watt bulb in response to the increase in demand created by the mandate, prices for 60 and 40 watt equivalent California compliant incandescents (which did not experience a similar increase in demand, see model 8 of table 3.5) also increased by similar amounts. These price increases appear to be unique to GE bulbs when other stores in California are studied, adding further suspicion that the mandates were to blame. Finally, if this price increase was attributable to the mandate, a similar price increase would have occurred in the rest of the country the next year when 100 watt incandescents were removed from that market, but that did not happen. Instead, the price increase could be a result of the fact that the manufacturer was shifting production away from the California compliant incandescents and retooling factories for the impending halogen incandescent roll out. At the very least, there is thin evidence that the reduction in quantity was due to the reduction in choice offered to consumers, but instead was likely attributable to the large increase in prices which may or may not have not been caused by the policy.

[14]as mentioned before,stores outside of California have very strong sales in the month of August, which is impossible to control for in a dynamic model like this without creating collinearity, but since this effect is consistent and similarly sized before and after the policy, it does not threaten the parallel trends assumption

2013, CA EISA 40/60w mandate, US 75w mandate

When the full EISA mandate was put in place in California at the beginning of 2013, the halogen incandescent replacements were finally made widely available, and all of the older Reveal or California compliant bulbs of any wattage equivalence were made illegal. A less efficient class of bulbs was completely replaced by a new, more efficient class. In this respect, the California mandate in 2013 looked like the 2008 mandate but on a bigger scale, without a loophole option (the Reveal bulbs), and with replacements (halogens) that had identical brightness as the old regular incandescents. I use this policy roll out to test the market power model (by looking for a decrease in the price of CFLs in response to an increase in the efficiency of incandescents). In addition, large scale stockpiling or reductions in overall demand would contradict the behavioral or indifference efficiency gap hypotheses. Unfortunately, stores outside of California were going through policy changes at the same time (being as they were only one year behind in the phase in of the mandate), so they do not provide the perfect control. Unsurprisingly, then, many more states register significant results in the placebo models reported in table 3.7.

For the third time, I see no evidence that the mandate increased sales of CFLs as some optimists had predicted. Model 5 in table 3.7 presents a fairly tight zero effect on CFL sales after the policy is fully implemented, and while the dynamic differences (figure 3.4, panel c) show strong seasonal variation of CFL consumption, neither year looks different than the other. Clearly, eliminating the bottom of the incandescent market entirely did not disrupt the overall incandescents market (now made up of more expensive and 30% more efficient halogens) enough to affect the CFL market. There is also no evidence of customers 'defecting' from the more expensive CFLs to the newly efficient incandescents. Panel d in figure 3.4 shows no price decrease in CFLs, which is strong evidence against a market power justification for the efficiency mandate.

Meanwhile, in the incandescent market, the price of a bulbs rose drastically with the switch to halogens (figure 3.4, panel b). Unlike the California compliant incandescents in 2008, halogen bulbs were immediately significantly more expensive than the bulbs they were replacing. Why, then, is there not a reduction in quantity of bulbs sold? In fact, when the triple difference is used to control for trends in the overall lighting market (model 4, table 3.7) a positive increase of bulbs sold in the neighborhood of 10% is seen as being caused by the mandate. Why, also, was there no evidence of stockpiling?

A number of logical explanations could explain an increase in demand (or at least, explain why no decrease in demand was seen as compared to national stores). The first explanation is that mandates were correcting for behavioral biases or rational indifference, and the mandates quietly corrected for this error, making everyone better off in the process. The lack of demand changes and stockpiling support this hypothesis, but there are some caveats.

Remembering that these difference models are comparing California to other stores that are also undergoing a policy change, it may be that the imposition of any type of mandate will reduce demand by a similar amount, thus California and the rest of the country were both 'treated'. This explanation is unsatisfying since the mandate being imposed on the rest

Table 3.7: 2013 California 40-60 Watt Efficiency Mandate

	Bulbs per Week per Store							
	All A-Shaped Bulbs			Triple Diff	CFLs	Incand	Non-Halogen	Halogen
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sale x CA	19.76***	22.80**	23.94***	−0.43	8.18***	11.58	−26.91**	38.36***
	(7.45)	(9.42)	(7.03)	(2.76)	(2.49)	(8.17)	(13.38)	(11.12)
Policy x CA	8.82	10.52*	19.54**	−19.68***	−1.12	9.94	−70.38***	80.19***
	(8.43)	(5.95)	(7.95)	(2.68)	(1.53)	(7.44)	(10.38)	(7.75)
Price			−75.33***					
			(5.42)					
A-Shape Bulbs				103.97***				
				(9.12)				
CA x A				−71.42***				
				(10.67)				
Sale x A				−2.14				
				(14.27)				
Policy x A				−26.51***				
				(7.68)				
Sale x CA x A				20.40***				
				(7.41)				
Policy x CA x A				28.64***				
				(7.54)				
Month x St FE	*No*	*Yes*	*No*	*No*	*No*	*No*	*No*	*No*
Week & Store FE	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
St Placebo Sig at 99%	5/11	–	3/11	5/11	1/11	5/11	6/11	4/11
Observations	174,791	174,791	174,791	349,582	174,754	174,791	174,819	174,819
R^2	0.75	0.77	0.76	0.75	0.75	0.72	0.71	0.74

Note: *p<0.1; **p<0.05; ***p<0.01
Standard Errors are clustered at the market and week level.
Pre period is wk ending 01-21-12 to 12-22-12. Sale is 12-29-12 to 01-26-13. Policy is 02-02-13 to 12-07-13

Table 3.8: December 2012 Week Store Avg

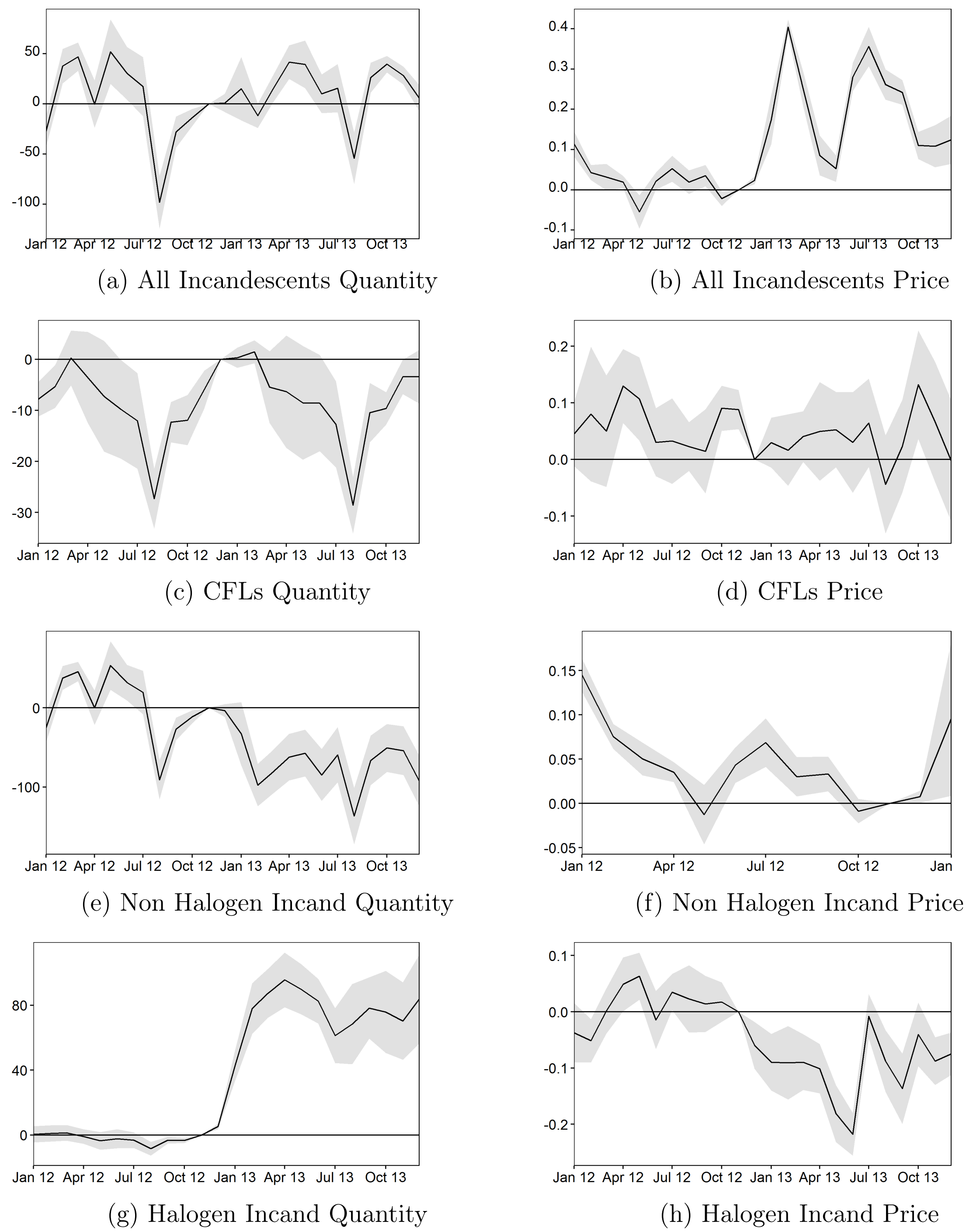
	California		Rest of USA	
	Bulbs	Price ($)	Bulbs	Price ($)
All Incandescents	238.5	1.33	332.1	0.94
All CFLs	85.8	4.6	71	4.62
Non-Halogen Incandescents	207	1.1	312.5	0.82
Halogen Incandescents	31.6	2.9	19.9	2.94

of the country covers so many fewer bulbs (75 watts is the least popular wattage) than the 40 to 60 watt equivalence mandate in California. It could also be that California faced a larger pent up demand in than the rest of the country when halogens were introduced in 2013. 100 watt incandescents had been unavailable in the state for two years by the time halogens were widely made available, while they had only been missing from shelves in the rest of the country for a year, and as I have shown, few consumers who were not already buying CFLs viewed them as worthy substitutes. If this explanation is true, California consumers did not so much embrace the new bulbs, but were desperate for any bulb at all that were not a CFL. This increase in demand could also be a story of wattage baselines. It may be that the consumers who switched to Reveal bulbs in 2008 viewed an increase in efficiency from the long standing standard with suspicion, but later did not lose any utility from a change from 57 watts to 43 watts. California consumers may have already gained more experience with new bulb technology and trusted them more, while consumers in the rest of the country reduced their consumption in response to the 75 watt ban. Finally, consumers may be hurt by the higher prices, but have no choice but to continue to buy bulbs or else be left in the dark.

Because the the rest of the country is no longer as clean of a control by this time, it is hard to use this evidence to say anything definitive about the direction of the effect of the full roll out of the mandate on the level of sales of bulbs, but the lack of evidence of any reduction in demand puts the findings of a reduction in demand from the previous two event studies into question.

If, however, we accept the theory that a lack of demand response meant consumers were as good or better off with the new, more efficient bulbs, why did it take a mandate to force suppliers to offer them? It is likely that behavioral biases would have limited the attraction of the new bulbs, and any one company who tried to manufacture halogens would have suffered from R&D and learning by doing spillovers. Halogen bulbs were clearly hard to produce, as they were not ready on a large scale until the third year of the mandate roll out. In addition, when they became available on a limited basis in 2011, the price was very high, but dropped steadily over the years, indicating that it would have been a risky investment for any company who wasn't required to undertake it (or, alternatively, who wasn't subsidized).

Figure 3.4: Differences in Store per Week Sales, California, 2012-2013

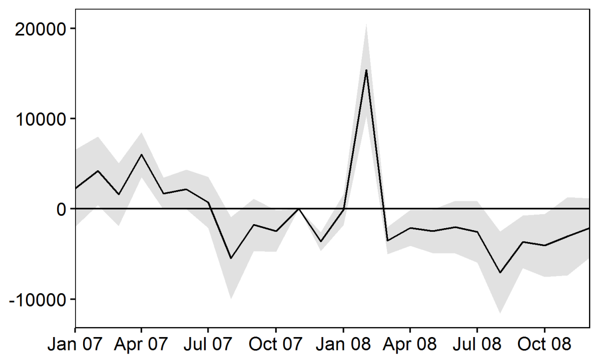


Note: Shaded areas are the 99% confidence intervals using standard errors clustered at the metro level. Results are from equation 3 using the category of bulb and either number of bulbs sold or average price on the left hand side. The omitted month for CFL is Dec 2007 while the omitted month for all the other panels is Nov 2007 (to avoid stockpiling). All incandescents are regular, CA compliant, Reveal, and halogen bulbs.

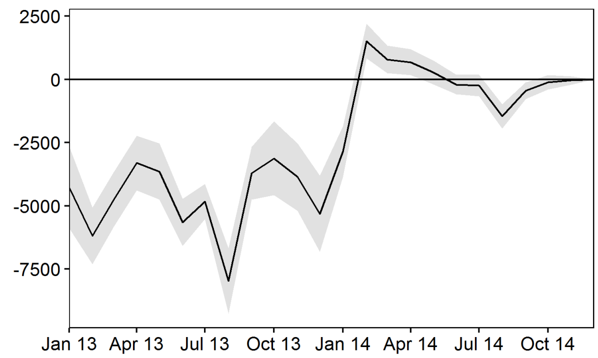
3.6 Policy Success and Welfare analysis

From a policy standpoint that is only concerned with improving the efficiency of the fleet of light bulbs, I find that the first California mandate in 2008 success was unclear, but the later policy mandate certainly increased efficiency. By recalculating equation three with the left hand side variable being the sum of watts from incandescent and CFL bulbs sold per store, I can analyze if there was a significant overall effect.

Figure 3.5: Differences in Total Watts Sold per Week, California



(a) 2008-2009 (5% Mandate)



(b) 2013-2014 (Rest of USA Adopts Full Mandate)

Note: Shaded areas are the 99% confidence intervals using standard errors clustered at the metro level. Results are from equation 3 using sum of watts from incandescent and CFL bulbs sold per store.

Panel a of figure 3.5 shows that the 5% mandate achieved marginally significant reductions. Most months, he reductions were in the 3,000 watts range (the average store in California in 2007 sold 35,000 watts worth of bulbs per month), but this reduction appears to follow a trend line established before the policy. In addition, the sale in February caused a massive spike in sales, meaning that it took about 6 months of the marginally significant reductions in order to 'even-out'. This change seems to be driven by reduced sales, and the reductions have very wide margins of error (which contract if reveal bulbs are excluded, suggesting that loophole purchases make the reductions lower and less certain). Stockpiling can clearly upend an efficiency mandate on light bulbs in the short term.

To estimate the policy outcome of the larger mandates, I analyze not the 2013 mandate, but the final mandate in the US in 2014 when regular incandescents were fully taken off the market everywhere (remember, they were off the market in California beginning in 2013). This allows me to see both treatment and control in a period where they have the same policy (in 2014) and gives me a time period (in 2013) where halogens were fully available in both places. Here, we see that before the mandate was applied nationwide, California stores clearly sold 5000 less watts per store and then afterwards, in 2014, the difference between the state's stores stayed steady. A similar analysis reveals that bulbs in California were 50 cents more expensive. Stores in California in 2013 sold an average of 10,300 watts per month, implying 33% savings, or 14 watts per bulb. Using the average marginal price of electricity of 15.5 watts per hour in 2013 in California, this improvement would pay off after about 204 hours of use, which seems like a reasonable payoff for even a semi-regularly used bulb. 204 hours is an upper bound for the policy, however, since 100 watt and 75 watt bulbs were already off the shelves in the US in 2013.

Whether this resulted in welfare gains across the board for all consumers is still difficult to prove. The lack of significant stockpiling and reductions in demand paired with large gains in efficiency at reasonably small price increases supports the success fo the policy, but examining the home panel data shows that stockpiling behavior existed at other stores not in the Nielsen scanner data. Still, it may be that stockpiling is not a bad thing. Stockpilers are probably not rationally inattentive, and while they may be making mistakes about the future, it is also possible that some households genuinely prefer the old bulbs or have rarely used sockets. In this case, stockpiling allows these households to avoid welfare losses form the mandate for a number of years. Stockpiling, then, is not necessarily a bad outcome.

3.7 Conclusion and Next Steps

From 2007, when the EISA was passed into law, to 2014, when its light bulb efficiency mandates were fully implemented, the number of light bulbs bought at the stores of a major national retailer fell dramatically and the price rose dramatically. Only part of the price increase was likely directly due to efficiency mandates (around 40 to 50 cents), and I find evidence that small mandates, like the one imposed by California in 2008, may have little to no positive price effect. Meanwhile, the evidence that mandates contributed to reductions in overall demand of bulbs is mixed. What is clear, however, is that these mandates did not nudge or force people to buy CFLs, as the supporters of the plans hoped. CLF prices did not decrease, which lessens the likelihood that the development of halogens was limited by market power and price discrimination. It is also clear that the mandates drastically increased the efficiency of the average bulb sold.

For California, and other energy hawkish states, there are other lessons to be learned. California proved it could impose its own efficiency mandates, independent of the rest of the country, but it did not go entirely smoothly. In order to preserve consumer choice, California had to exempt specialty bulbs, which resulted in leakage (when the full EISA mandates were imposed nationwide, manufacturers, likely responding to a larger pool of consumers, came out with halogen versions of the premium bulbs). Enforcement was also a significant problem. As figure 3.1 shows, the national retailer reintroduced the illegal regular incandescents in California stores for the first half of 2009, and the full Nielsen scanner data show that other retailers were slower to adjust to the California rules in 2008 and in 2011-2013 than they were in the rest of the country. Cheating by retailers is impossible under a national mandate if all manufacturers simply switch their production. If manufacturers do not adopt California's mandated changes to the nationwide market, as they did not in 2008, it exposes California consumers to a smaller supply chain with unexpected and unexplained price changes as likely happened in late 2011. Finally, going it alone on mandates exposes a states consumers to the risk that a replacement product will not be ready for them immediately. In 2008, the national retailer waited a number of months to switch to the new bulbs because California compliant bulb production did not start until March, and if you were a consumer in California who wanted to buy a 100 watt equivalent incandescent bulb in January 2011, you would have to wait two years.

Many of the unanswered questions could be attempted to be answered by a detailed analysis of the household panel. Sales of light bulbs, while a relatively frequently purchased appliance, is still bought fairly irregularly in this dataset, but a discrete choice model would possibly be a valuable extension to this literature. As mentioned before, stockpiling was clearly more prevalent in the household data, suggesting that this behavior may be dependent on the type of store. Summary statistics of the household data do support my other conclusions, however. CFL consumption dose not increase, and their prices do not go down after the mandate.

Bibliography

Allcott, Hunt, and Michael Greenstone. 2012. "Is there an energy efficiency gap?" *The Journal of Economic Perspectives* 26 (1): 3–28.

Barbose, Galen L, Charles A Goldman, Ian M Hoffman, and Megan Billingsley. 2013. "The future of utility customer-funded energy efficiency programs in the USA: projected spending and savings to 2025." *Energy Efficiency* 6 (3): 475–493.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How much should we trust differences-in-differences estimates?" *The Quarterly journal of economics* 119 (1): 249–275.

CEC. 2004. "Committee Hearing, Oct 13." Accessed February 15, 2017. `http://www.energy.ca.gov/appliances/archive/2004rulemaking/documents/2004-10-13_TRANSCRIPT.PDF`.

EPA. 2005. "Energy Efficiency in the Clean Power Plan." Accessed February 15, 2017. `https://www.epa.gov/cleanpowerplan/fact-sheet-energy-efficiency-clean-power-plan`.

———. 2011. "EISA Frequntly Asked Questions." Accessed February 15, 2017. `https://www.energystar.gov/ia/products/lighting/cfls/downloads/EISA_Backgrounder_FINAL_4-11_EPA.pdf?6bd2-3775`.

Fischer, Carolyn. 2005. "On the importance of the supply side in demand-side management." *Energy Economics* 27 (1): 165–180.

Fowlie, Meredith, Michael Greenstone, and Catherine Wolfram. 2015. *Do energy efficiency investments deliver? Evidence from the weatherization assistance program.* Technical report. National Bureau of Economic Research.

Gellings, Clark W, and JH Chamberlin. 1988. "Demand-side management." *Energy Efficiency and Renewable Energy Handbook, Second Edition edited by D. Yogi Goswami, Frank Kreith (Chapter 15):* 289–310.

Gerarden, Todd, Richard G Newell, and Robert N Stavins. 2015. "Deconstructing the energy-efficiency gap: conceptual frameworks and evidence." *The American Economic Review* 105 (5): 183–186.

Gillingham, Kenneth, Richard G Newell, and Karen Palmer. 2009. "Energy efficiency economics and policy." *Annu. Rev. Resour. Econ.* 1 (1): 597–620.

Gillingham, Kenneth, and Karen Palmer. 2014. "Bridging the energy efficiency gap: Policy insights from economic theory and empirical evidence." *Review of Environmental Economics and Policy* 8 (1): 18–38.

Granade, Hannah Choi, Jon Creyts, Anton Derkach, Philip Farese, and Scott Nyquist. 2009. "Unlocking energy efficiency in the US economy."

Gromet, Dena M, Howard Kunreuther, and Richard P Larrick. 2013. "Political ideology affects energy-efficiency attitudes and choices." *Proceedings of the National Academy of Sciences* 110 (23): 9314–9319.

Ito, Koichiro, and James M Sallee. 2014. *The economics of attribute-based regulation: theory and evidence from fuel-economy standards.* Technical report. National Bureau of Economic Research.

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

Sallee, James M. 2014. "Rational inattention and energy efficiency." *The Journal of Law and Economics* 57 (3): 781–820.

Shogren, Jason F, and Laura O Taylor. 2008. "On behavioral-environmental economics." *Review of Environmental Economics and Policy* 2 (1): 26–44.

Smith, Rebecca. 2008. "Utilities Amp Up Push To Slash Energy Use." *The Wall Street Journal* (January 9).