

**DEMAND EXTERNALITIES FROM CO-LOCATION**

**By**

**Boudhayan Sen, Jiwoong Shin and K. Sudhir**

**February 2012**

**COWLES FOUNDATION DISCUSSION PAPER NO. 1850**



**COWLES FOUNDATION FOR RESEARCH IN ECONOMICS  
YALE UNIVERSITY  
Box 208281  
New Haven, Connecticut 06520-8281**

**<http://cowles.econ.yale.edu/>**

# Demand Externalities from Co-Location

Boudhayan Sen, Jiwoong Shin, and K. Sudhir<sup>†</sup>

*Yale University*

First Draft: June 2011

This Draft: February 2012

---

<sup>†</sup>Boudhayan Sen is a PhD candidate in Marketing (e-mail: boudhayan.sen@yale.edu), Jiwoong Shin is an Associate Professor of Marketing (jiwoong.shin@yale.edu), and K. Sudhir is James L. Frank Professor of Private Enterprise and Management (k.sudhir@yale.edu) at School of Management, Yale University, 135 Prospect Street, New Haven, CT 06520. We thank Joseph Altonji, Todd Elder, Minha Hwang, Harikesh Nair, Sangwoo Shin, and seminar participants at McGill and Yale University for their helpful comments.

# Demand Externalities from Co-Location

## Abstract

We illustrate an approach to measure demand externalities from co-location by estimating household level changes in grocery spending at a supermarket among households that also buy gas at a co-located gas station, relative to those who do not. Controlling for observable and unobserved selection in the use of gas station, we find significant demand externalities; on average a household that buys gas has 7.7% to 9.3% increase in spending on groceries. Accounting for differences in gross margins, the profit from the grocery spillovers is 130% to 150% the profit from gasoline sales. The spillovers are moderated by store loyalty, with the gas station serving to cement the loyalty of store-loyal households. The grocery spillover effects are significant for traditional grocery products, but 23% larger for convenience stores. Thus co-location of a new category impacts both inter-format competition with respect to convenience stores (selling the new category) and intra-format competition with respect to other supermarkets (selling the existing categories).

**Key Words:** Revenue economies of scope, demand externalities, one stop shopping, co-location, selection, retail industry.

# 1. Introduction

The introduction of new, co-located products and services is a ubiquitous phenomenon among retail firms. Wal-Mart, for instance, now offers a greatly expanded set of products and services at their stores— from vision centers (eye care) and clinics to auto maintenance services (e.g. tire replacements). Supermarkets have also followed this trend, incorporating coffee shops, DVD rental kiosks and gas stations (among other products) into their store layouts. An important managerial question that arises in launching new services is: Does co-locating such new services in the store lead to demand externalities in the form of increased sales of the existing product assortment?

The issue of demand externality from spatial co-location is of broader interest than for retailers. Mall developers, need to evaluate inter-store demand externalities in evaluating store mix in malls. In practice, anchor stores often serve to create foot traffic that provides positive demand externalities for smaller stores and hence are offered rental discounts (Benjamin et al. 1992; Gould et al. 2005). Smaller stores are also believed to collectively create positive demand externalities for the larger anchor stores as well. Faced with dilapidated downtowns, many local governments and civic organizations have begun to encourage big-box store openings in downtown areas (Philips 2010), hoping to generate demand externalities for smaller stores through increased foot traffic. Gould et al. (2005) speculate that one potentially important reason for the decline of the central business districts across the United States was their inability to price the externalities of anchor stores, leading to the flight of large anchor stores from downtown areas to malls.

Despite its significance, there is little work on establishing the existence of demand externalities from spatial co-location. The limited research typically uses indirect approaches to quantify demand externalities through the supply side choices of firms. For instance, in the entry games literature, externalities are quantified from the decisions of major retailer to enter (or not enter) a mall (Vitorino 2010), or from the entry and format choice decisions of retailers operating within a retail zone (Datta and Sudhir 2010). Gould et al. (2005) estimate externalities among retailers in a mall based on the systematic differences in rental rates offered by mall operators to anchor stores versus smaller stores. In all of these papers, externalities are quantified under the assumption that firms already have knowledge of the magnitude of externalities and make optimal decisions given this information.

However, firms may not necessarily have accurate knowledge of the level of externalities and casual observation shows that firms often revise their decisions after initially increasing scope. For instance, Stop & Shop, a northeastern supermarket, shut down Dunkin Donuts coffee shops at several of its locations, presumably due to limited demand externalities relative to the opportunity cost of in-store space. In another instance, First Tennessee bank closed several of its branches that were located within grocery stores, citing that the bank “found customers at its grocery locations were not using the array of services there...” (Bailey 2009). These instances suggest that managerial intuition about the existence of favorable demand externalities may not be accurate, and hence there is the need for an alternative approach that is not dependent on managerial decision accuracy.<sup>1</sup>

We therefore consider an alternative approach that measures demand externality from the new product on the larger existing assortment using directly observable, consumer-level changes in grocery spend and the number of trips to the store. In particular, we illustrate this direct approach in the context of a supermarket that opens a gas station in its parking lot. The phenomenon of grocery stores with co-located gas stations has become relatively widespread in recent years and hence the understanding of the demand externalities from introducing gas stations is of substantive interest in its own right (Goic et. al. 2010). The approach also has the managerial advantage that any retailer can replicate the analysis using data that is available at the store.

We use our measurement approach to address the following questions: Does co-locating the secondary small gas business help expand the primary business of groceries? And if yes, what is the magnitude of this externality? Further, which customers contribute to the grocery externality—loyal shoppers or casual shoppers? From which competitor does the store derive the grocery externality—other supermarkets or convenient stores associated with gas stations? Thus, direct measurements can (1) not only help managers and firms obtain better insights for decisions

---

<sup>1</sup> Contrary to our lay belief, research on economies of scope in the banking (between depository and loan services) and insurance (between life and casualty insurance) industries has failed to find evidence of demand externalities (Berger et al. 1996, 2000; Cummins et al. 2010). In a different context, a recent McKinsey study finds that 70 percent of mergers failed to achieve expected revenue synergies (Christofferson 2004, Frieswick 2005). In a similar vein, the Economist notes remarks that “activities that are chronically unprofitable are carried on [by European banks] in the belief that they help a bank to keep its customers” (Economist 1993, Klemperer and Padilla 1997). Gaffen (2010) discusses Citigroup’s lack of success in becoming a one-stop financial supermarket for its customers.

on the scope of assortments, (2) but also obtain greater insights on the changes in target customer segments and competitive market structure.

As there are no obvious complementarities in consumption between groceries and gasoline, demand externalities among existing customers – if they exist – are likely to be driven by the benefits of one-stop shopping, or an effectively lower travel cost to the store (Bell et al. 1998, Gauri et al. 2008, Wernerfelt 1994). We test the one-stop shopping/low travel cost hypothesis by examining whether increased spending among gas-buyers accompanied by an increase in trips to the store?<sup>2</sup>

Even with detailed household level data, before and after opening the gas station, establishing the existence and empirically measuring magnitude of demand externalities is challenging for several reasons. First, households choose whether and when to become gas-buyers. Since “buying gas” is not an exogenous treatment, and households self-select into this treatment condition, we need to address the effects of observable and unobservable selection effect to obtain the true effect of demand externality from co-location. Second, the gas station may have been opened at a particularly high spending season for groceries, and hence the higher spend among gas buyers may not be permanent and give a biased estimate of spillover benefits. Finally, increases in spending among households may be driven by unobserved events around the opening of the gas station, such as increased promotional activity or unreported improvements to store operations. We need to rule out such “fortuitous” timing and endogenous promotion explanations.

Our basic empirical strategy is as follows: We first compare the within-household change in grocery spend and trips “*before and after*” the introduction of the gas station between households *that buy gas at the new gas station relative to those who do not*. This provides us a preliminary estimate of demand externalities from co-locating the gas station at the grocery store. To exclude household-specific explanations about the time they started using the gas station in the before-

---

<sup>2</sup> In addition, conversations with managers at the retailer indicated that there was no effort to position gasoline as a loss-leader or undercut the prices of nearby gas stations. Households were given a constant incentive of a 5-cent per gallon discount as an incentive for identifying themselves by using their loyalty card when purchasing gasoline. Managers indicated that the gasoline prices even after this discount were comparable with prices at nearby gas stations. Related research shows that households shift their spending towards lower priced products (Gicheva et. al. 2010) and towards supercenters and warehouse clubs in response to increasing gas prices (Ma et. al. 2011). Since the gas station in our data opened during a period of gradually falling gasoline prices, it seems unlikely that the treatment store location becomes more attractive because of favorable gasoline prices.

after analysis, we restrict our initial analysis to a short window around the time of the gas station opening (6-weeks before and after the gas station opening) and simply compare changes in spending at the household level on a common pre and post period, based on when the gas station was opened, rather than when the household started purchasing gas.

After establishing the preliminary estimates of demand externality from co-location, we assess the robustness of these estimates to selection on observables using propensity matching (DiPrete and Gangl 2004, Leuven and Sianesi 2003). Furthermore, to address the possibility of additional hidden bias due to selection on unobservables, we adapt an approach developed by Altonji, Elder and Taber (2002, 2005) to provide a conservative lower bound on the size of the externality. Based on these approaches, we find the evidence of demand externalities with increases over the short-term of 7.7%-9.3% in grocery spend among gas-buyer households. These increases in spending are accompanied by a 14%-15% increase in the number of trips made to the store, indicating that increased spending likely occurs due to the lower travel costs associated with one-stop shopping. These spillovers are economically and managerially significant. Given our estimates, grocery spillovers are about 18-22% of gas revenues. Even more impressively, since groceries have much higher margins than gas, the gross profits from grocery spillovers are 130-150% of gross profits from gasoline sales.

We then assess other concerns of potential endogeneity about the timing of the gas station opening, and effect of unobservable events such as promotional activity around the time of the gas station opening. We show that the timing of the gas station opening does not appear to drive our findings about demand externalities, and that the effect of promotional activity (or other unobserved events) around the time of the gas station opening do not account for the increase in spending among gas buyers. Moreover, we find that gas-buyers tend to increase their grocery spending by similar amounts regardless of the timing of when they first use the gas station. Finally, although the incremental spending does diminish over time, incremental grocery spending by gas-buyers remains positive and significant over a long time-frame.

We then seek a deeper understanding of the mechanism by investigating some moderators of demand externalities in our context. Specifically, we study the moderating effect of customer loyalty to the supermarket on the extent of grocery spillovers. We find that the households with relatively high loyalty show the greatest response in grocery spend (both absolute and relative) to the increased scope of the firm, which suggests that the gas station serves to consolidate spending

among households that are already loyal to the store. Finally, we test how the introduction of gas station impacts competitive market structure: Does it increase competition with respect to other supermarkets or other convenience stores, typically associated with gas stations? We find that the additional grocery dollars come at the expense of both traditional grocery competitors as well as convenience stores, but there is a greater percentage of convenience store purchases that are shifted to the supermarket introducing the gas station.

Overall, this paper contributes along several dimensions. As a measurement approach, we offer a direct measurement approach for measuring demand externalities using household-level data that are easily available to retailers. Because it does not require assumptions of managerial optimal behavior, the approach can guide managerial decision making at each store level. Qualitatively, our results demonstrate that even a secondary service can have a large spillover impact on the profits from the primary business, and serves to cement the loyalty of core users of the primary product. Methodologically, we adapt the Altonji, Elder and Taber (2002, 2005) bivariate probit framework for placing bounds on the combined effects of selection on observables and unobservables to allow for a continuous outcome variable such as grocery spend. Zhu and Sun (2012) apply our approach in their study of content creation by bloggers.

The structure of this article is as follows: Section 2 describes the data and Section 3 presents the main results on the extent of demand externalities from increased scope, controlling for observed and unobserved selection. Section 4 addresses concerns relating to potential endogeneity about the timing of the gas station opening, confounding effects arising from concomitant promotional activity with the opening of gas station, and whether the spillover effects of gas station last over a longer time-frame. Section 5 presents analysis of moderators of demand externalities and we conclude in Section 6.

## 2. Data

We obtain household level transaction data at the category level for two geographically distant stores of a major supermarket chain within Connecticut over a 52 week period between June 2005 and June 2006. One store opened a gas station in its parking lot in September 2005, (14th week into the data period) while the other store opened a gas station in 2001, well before our data period. We will refer to the first store as the “treatment” store and the other as “control



store with gas,” respectively. Further, we will call the weeks prior to the introduction of the gas station at the treatment store as “pre-gas” and the weeks after introduction as “post-gas” period. In the analysis, we will use alternative window lengths for the pre and post-gas periods as appropriate, but pre- and post are anchored at the week in which the gas station was opened at the treatment store.

The retailer also provided internal estimates of household “share of wallet (SOW),” which are helpful to answer questions about the moderating impact of loyalty on demand externalities. These internal estimates of SOW are based on estimates of potential grocery spend for a household created by a third party firm. This firm uses a proprietary algorithm that accounts for the household’s geo-demographic characteristics to develop the potential estimate. It is important to note that the SOW metric was constructed prior to the time-period of our data, and it is not influenced by the changes in spending induced by the gas station.

The grocery spend amounts are categorized into 593 categories, defined internally by the retailer. To understand how the introduction of the gas station affects convenience stores and supermarkets differentially, we classified the 593 categories into 464 “traditional grocery” categories and 129 “convenience categories.”<sup>3</sup> As examples, cigarettes are a “convenience” category, while fresh vegetables are a “traditional” category.

### *Descriptive Statistics*

Table 1a summarizes average grocery spend and gas station use across the two types of stores, before and after the gas station was opened at the treatment store. Grocery spend per week per household increases from the pre to the post period only at the treatment store. Further, gas use penetration at the treatment store increases over the post-gas period and becomes comparable to the control with gas store within 14 weeks.

\*\*\* Table 1 \*\*\*

---

<sup>3</sup> We surveyed four grocery shoppers (who are primary shoppers in their households) to classify categories into two mutually exclusive groups: “convenience” products and “traditional grocery”. The classification across these respondents was highly consistent (Cronbach’s  $\alpha$  of 0.87). When there were occasional disagreements, we used the classification chosen by the majority of respondents. The exhaustive classification is available from the authors upon request.

Next, we compare the average spend and trips between gas buyers and non-gas buyers at the treatment store in Table 1b. In order to isolate the change in spend to the effect of the gas station, the analysis is around a twelve-week period around the opening of the gas station – a six week pre-period and six week post-period.<sup>4</sup> Overall, households spend an average of \$41 per week on groceries, and this stays stable across households between pre- and post-periods. But there are clear differences among those who buy gas in the six week post period and those who do not. The average weekly grocery spend for the non-gas buying household in the post-period falls by 6% relative to the pre-period ( $p < 0.01$ ); while the spend increases by 7% for households that purchased goods ( $p < 0.01$ ). Further, households who purchase gas tend to a priori spend more at the store (with average weekly spending for grocery in the pre-periods being \$50.78 for gas-buyers and \$35.04 for non-buyers, respectively). Households that purchased gas in the post-period spent on average \$13.80 for gasoline per week.

In terms of trips, as for spend, there is no statistically significant difference in the number of trips between pre- and post-periods, with overall household trips to the store at 0.86 and 0.85 trips per week in the pre- and the post-periods, respectively. But there is a significant difference between the gas-buyers and non-buyers. For households who did not buy gas, there was a decrease in trips between the pre- and post- periods from 0.73 to 0.66 ( $p < 0.01$ ). Among households that bought gas, average trip per week indeed increases from 1.07 to 1.13 ( $p < 0.01$ ). Finally, since there are no consumption complementarities between the grocery products and gasoline, the percentage increase in grocery spending and trips among households that bought gas provides prima-facie evidence for the demand externalities stemming from the one-stop convenience created by the gas station. But the differences in grocery spend and number of trips in the pre-period suggests that there are “intrinsic” differences between gas buyers and non-buyers—implying that potential selection issues need to be addressed.

---

<sup>4</sup> The rationale to use small time window is similar to arguments in regression discontinuity approaches (for example, Busse et al. 2006), in which local average treatment effects are estimated from changes within a short window around a discontinuity. The shorter the time period, the stronger the causal claim that the differences in behavior are driven by the introduction of the gas station and not due to other household-side unobservables that evolve over time. We will assess sensitivity to the time window and issues of selection in subsequent sections.

### 3. Analysis and Results

Our analysis strategy is as follows: We begin with a cross-sectional regression of difference in grocery spend and trips (in absolute and percentage terms) before and after the gas station is opened in order to test whether households that purchased gas at the newly introduced gas station buy more groceries than those who don't. We interpret this incremental spend and trips among households that purchased gas as a measure of the demand externality. As explained earlier, to minimize the impact of household unobservables that evolve over time, we perform the initial analysis with a short time window—6 pre and 6 post weeks. We assess the robustness of the results to (1) alternatively specified models—for example, a panel regression of grocery spend with household and weekly fixed effects and a Poisson regression for the number of trips and (2) using a longer time window—14 pre and 14 post weeks.

In our analysis, we treat households that do not purchase gas as the “control.” We assess the reasonableness of this group as “control” group using data from other (control) store with gas. We rule out a competing explanation for our results: the difference in spend between gas and non-gas buyers may not be due to incremental spend by gas buyers, but a decline in spend by non-gas buyers due to negative impact of gas station to non-gas-buyers (e.g., due to higher store traffic). Finally, we demonstrate robustness to concerns about observable and unobservable selection effects, because gas buyers self-select to purchase gas.

#### 3.1 Demand Externality from Co-location

##### *Impact of Gas Station Co-location on Grocery Spend*

To estimate demand externalities from co-location, we compare two differences in household behavior at the treatment store. The first difference is a temporal “before-after” comparison based on differences in grocery purchase behavior before and after the gas station was opened. The second difference is cross-sectional between households who buy gas versus those who do not after the gas station is opened.

Let the grocery spend of household  $i$  at the treatment store before and after the gas station opening be given by  $Groc_i^{pre}$  and  $Groc_i^{post}$ . Denote the difference in grocery spending of household  $i$  before and after the introduction of gas station, as  $\Delta Groc_i = Groc_i^{post} - Groc_i^{pre}$ . To estimate the

difference in spending between households that purchase gas versus those that do not, we estimate the following model

$$\Delta Groc_i = Groc_i^{post} - Groc_i^{pre} = \alpha_0 + \beta \cdot 1\{Gas\_Buy\}_i + \varepsilon_i, \quad (1)$$

where  $1\{Gas\_Buy\}_i$  is an indicator variable for whether household  $i$  purchases gas in the post-period, and  $\varepsilon_i$  is the household  $i$  specific error. The coefficient  $\beta$  captures the difference in difference and is the parameter of interest. The estimates are reported in Table 2, Column 1.<sup>5</sup>

\*\*\* Table 2 \*\*\*

On average, gas-buyers spend \$5.78 more on groceries per week, relative to non-gas buyers after introduction of the gas station. This difference is statistically significant at the 1 percent level, providing evidence of demand externalities on grocery spend from the co-location of the gas station.

Since households differ substantially in their pre-period purchases, and the gas station is likely to have a proportionate impact on household grocery spend, we next estimate the percentage change in spend per household after the introduction of the gas station by estimating the following specification:

$$\Delta\%Groc_i = \frac{Groc_i^{post} - Groc_i^{pre}}{Groc_i^{pre}} = \alpha_0 + \beta \cdot 1\{Gas\_Buy\}_i + \varepsilon_i. \quad (2)$$

Column 2 of Table 2 reports the results. We find that gas buyers spend 13% more on groceries relative to non-gas buyers after the gas station is introduced. This effect is consistent with the descriptive statistics reported in Table 1, and reflects significant demand externalities on grocery spend from co-locating the gas station.

### *Impact of Gas Station Co-location on Grocery Trips*

---

<sup>5</sup> The metric of grocery spending in this analysis and for the rest of the paper is gross sales – i.e. revenues before accounting for discounts on specific grocery products. For robustness, we conducted this analysis using net sales (i.e. including the effect of grocery product-specific discounts); the results do not change.

Co-locating the gas station may cause increase in grocery spend for two reasons. First, trips that might previously have been made to gas stations, along with incidental spending at gas-station convenience stores, are now made to the focal grocery store. Second, after the introduction of the gas station, households may shift their regular grocery trips away from other grocery stores that do not sell gas. Instead, these trips are made to the focal supermarket, thereby allowing customers to avoid the additional trip required for gasoline. Either of these hypotheses requires that the number of trips to the focal grocery store should increase after the introduction of a gas station.

Let the number of trips of household  $i$  before and after the gas station opening be given by  $Trip_i^{pre}$  and  $Trip_i^{post}$ . Denote the difference in grocery trips of household  $i$  before and after the introduction of gas station, as  $\Delta Trip_i = Trip_i^{post} - Trip_i^{pre}$ . Then to estimate the difference in trips between households that purchase gas versus those that do not, we estimate the analogous model for changes in number of trips and percentage change in number of trips below:

$$\begin{aligned} \Delta Trip_i &= Trip_i^{post} - Trip_i^{pre} = \alpha_0 + \beta \cdot 1\{Gas\_Buy\}_i + \varepsilon_i, \\ \Delta\%Trip_i &= \frac{Trip_i^{post} - Trip_i^{pre}}{Trip_i^{pre}} = \alpha_0 + \beta \cdot 1\{Gas\_Buy\}_i + \varepsilon_i. \end{aligned} \tag{3}$$

The results are reported in Columns 3 and 4 of Table 2. The number of trips increases by 0.13 trips per week among households using the gas station, relative to those who do not; in percentage terms, trips increase by 15% (statistically significant at the 1 percent level). From these results, we conclude that co-location of the groceries and gasoline leads to consolidation of trips.

Finally, relative changes in grocery spend and trips are robust to alternative time-windows. We present results from a 28-week period (14 weeks, before and after), which is the largest available window in the data set. These results continue to be in line with our original results (Columns 5 and 6 in Table 2).

### *Panel Regressions*

In the above regressions, we aggregate spend and trips by a household over a certain pre and post periods and compare differences across the pre and post-period. An alternative approach that allows for week to week differences in the level of spend would be a panel specification with both

time and household specific fixed effects. We therefore estimate the following panel specification, incorporating household fixed effect ( $\alpha_i$ ) and week fixed effect ( $\tau_t$ ):<sup>6</sup>

$$\begin{aligned} \text{Groc}_{it} &= \alpha_i + \tau_t + \beta \cdot 1\{\text{Gas\_Buy}\}_i \cdot 1\{\text{Gas\_Station\_Open}\}_t + \varepsilon_{it} \\ \text{Trip}_{it} &= \alpha_i + \tau_t + \beta \cdot 1\{\text{Gas\_Buy}\}_i \cdot 1\{\text{Gas\_Station\_Open}\}_t + \varepsilon_{it} \end{aligned} \tag{4}$$

The interaction effect is the incremental effect of ‘‘Gas Station’’ on groceries. For ease of comparison, we use the same 12 week period as in the original specification. The results, reported in Columns 7 and 8 of Table 2 show that the magnitude of the effects are similar to what we obtain with the previous specification.

We also consider alternative functional forms. Using the natural log of grocery spend as the dependent variable, we find a positive and significant effect of becoming a gas-buyer (Column 9 of Table 2). As trips are a count variable, we also estimate a Poisson specification. The results in Column 10 of Table 2 also show similar effects.

### 3.2 Ruling out Other Explanations

An alternative explanation for our results above is that there is a rising trend in the difference in spends between gas and non-gas buyers independent of the introduction of the gas station. We test for such a trend by comparing the difference in spend between gas buyers and non-gas buyers, 1-6 weeks before gas station opens and 7-12 weeks before the gas station opened. Table 3, Column 1 shows no evidence of such a trend before the introduction of the gas station.

\*\*\* Table 3 \*\*\*

A second concern is that our estimated difference may not be due to gas-buyers, but non-gas buyers buying less. A plausible reason for non-gas buyers buying less is if the introduction of the gas station and the potentially higher traffic at the store made the shopping experience worse to households that are not interested in buying gas at the store. Given that grocery spend for non-gas buyers fell by 6.3% (see Table 1) after the gas station introduction, it is important to rule out this alternative explanation. We therefore compare the differences in spend before and after the

---

<sup>6</sup> Note that we do not estimate the ‘‘Gas\_Buy’’ and ‘‘Gas\_Station\_Open’’ main effects, because those are absorbed in the household and weekly fixed effects respectively.

gas station is introduced for non-gas buyers at the “treatment store” and the “control store with gas,” where the gas station was introduced in 2001. Table 3, Column 2, reports the results of a regression with the difference in spend before and after the gas station opened as the dependent variable and a treatment store dummy variable. If the negative impact of the gas station drives the results, the decrease in spend among non-gas buyers should be greater at the treatment store i.e., the treatment store dummy should be negative. However, the estimate is positive; suggesting that the negative impact of gas station on non-gas buyers is not driving our results.<sup>7</sup>

### 3.3 Selection

As discussed in the introduction, since households self-select into the “treatment” condition of whether they “buy gas” or not, it is possible that the difference in grocery spend/trips before and after gas station opens for gas buyers and non-gas buyers are subject to selection effects. Specifically, if changes in spend between the pre- and post-period of introducing the gas station affect gas and non-gas buyers differentially, the estimated incremental spend that we attribute to the opening of gas station will still be biased due to selection. It should be noted that to the extent that we compare percentage differences in spending before and after the gas station as the outcome variable, selection effects may be minimized in our earlier analysis, because we do allow for proportional differences in spend and trips between the control and treatment groups. Nevertheless, we assess the robustness of our results to such selection effects in this section. We first consider selection on observables and then selection on unobservables.

#### *Selection on Observables*

We use propensity matching to control for selection on observables. The basic idea is to ensure that the comparisons are made between households in the control and treatment groups, who are as similar as possible on observable characteristics in terms of their propensity to be in the treatment group (DiPrete and Gangl 2004, Leuven and Sianesi 2003).

---

<sup>7</sup> We also compared the time trend in spend among non-gas buyers at the treatment and control store prior to the opening of the gas station. We compared grocery spend, 1-6 weeks and 7-12 weeks before gas station opened. Though there was a downward trend for both stores, there was no difference in trend at the two stores. We therefore conclude that the reduction in grocery spending among non-gas buyers is a general time-trend, and not due to the negative impact of the new gas station.

Note from Table 1 that gas-buyers spend more on average on groceries per week relative to non-gas buyers – \$50.78 as opposed to \$35.04 – even before the gas station opened. Also gas buyers make more trips to the store: non-buyers average 0.73 trips per week, while gas-buyers average 1.07 trips per week before the gas station opened. Therefore we use pre-gas grocery spend and trips to match households on the propensity score. Our estimate of spillover reported in Row 2 of Table 4a shows that selection is an important issue. The incremental percentage change in spend based on propensity matching falls from 13.3% to 9.6%. Interestingly, the percentage change in trips is almost the same as in the non-propensity-scored estimates (14.7% vs. 14.6%).

\*\*\* Table 4-a \*\*\*

We also consider other variables that are of potential interest to match. Rather than control for spend at the store we decompose spend into its constituent components: “potential spend” and “share of wallet.”<sup>8</sup> These results are reported in row 3 of Table 4a. We find that gas-buyers increase their weekly grocery spend and trips by \$6.85 and 0.16, respectively, relative to non-buyers. In relative terms, the percentage differences in grocery spend and trips are 9.3% and 14.1%, respectively. Thus, even after controlling for selection on observables, the demand externalities on grocery spend and trips are substantial.

### *Selection on Unobservables*

We next address potential selection on unobservables by adapting an approach developed by Altonji, Elder and Taber (2002, 2005; henceforth AET) in the labor and econometrics literature.<sup>9</sup> Here, the primary identification problem arises because of a lack of plausible instruments: i.e. we don’t have a covariate that affects grocery spend (or trips made) solely through the decision to use the gas station. The AET approach is based on the assumption that selection on unobservables is less than or equal to selection on observables,<sup>10</sup> which allows the specification of a restricted correlation term between unobservables in both the outcome and selection equations (discussed

---

<sup>8</sup> Note that Spend=Potential Spend\*SOW. Potential Spend is calculated using the observed spending in the pre-period divided by SOW. To avoid potential multicollinearity, we use log(potential spend) for the propensity matching control.

<sup>9</sup> In a previous version of the paper we controlled for selection on unobservables using the Rosenbaum(2002) Bounds approach. The results from that approach also provided convergent validity to the findings using our extension of the AET approach. These results are available from the authors upon request.

<sup>10</sup> This assumption is reasonable if the observable variables indeed are good predictors of the dependent variable.



below). Not having to separately identify the correlation term allows the overall model to be identified.

Similar to AET, we specify the outcome and selection equation as follows:<sup>11</sup>

$$\begin{aligned}\Delta\%Groc_i &= \alpha \cdot 1\{Gas\_Buy\}_i + X_i'\gamma + \varepsilon_i, \\ Gas\_Buy_i^* &= X_i'\beta + u_i, \\ \text{where } \begin{pmatrix} \varepsilon_i \\ u_i \end{pmatrix} &\sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2 & \rho\sigma \\ \rho\sigma & 1 \end{pmatrix}\right)\end{aligned}\tag{5}$$

Here, we model the selection using the latent variable,  $Gas\_Buy^*$ , and the outcome of interest using the same covariates ( $X$ , which is comprised of the same three variables used for propensity scoring:  $log(potential)$ ,  $SOW$  and  $trips$  in the pre-period).

Unlike the traditional approach to selection problems, which typically utilize one or more exclusion restrictions, AET do not require exclusion restrictions, but achieve identification through the use of an additional restriction on  $\rho$  as follows. We begin by projecting the latent variable that represents the choice to buy gas onto the other components of the outcome equation. We can then represent the projection as

$$\text{Proj}(Gas\_Buy_i^* | X_i'\gamma, \varepsilon_i) = \phi_0 + \phi_{X'\gamma} X_i'\gamma + \phi_\varepsilon \varepsilon_i\tag{6}$$

Restricting the effect of unobservables to be relatively the same as the effect of observables is then equivalent to specifying that  $\phi_{X'\gamma} = \phi_\varepsilon$ .<sup>12</sup> It is easy to see that the correlation between the residuals in the selection and outcome equations lie in the range.

$$0 \leq \rho \leq \sigma \left[ \frac{Cov(X'\beta, X'\gamma)}{Var(X'\gamma)} \right]\tag{7}$$

---

<sup>11</sup> The key difference with respect to AET is that they develop the approach for bivariate probit model; in contrast, our outcome variable is continuous. As their outcome variable is discrete, they normalize the variance of  $\varepsilon_i$ ,  $\sigma$ , to 1.

<sup>12</sup> Note that the restriction  $\phi_{X'\gamma} = \phi_\varepsilon$  is more realistic than the standard OLS approach, which would require that  $Cov(Gas\_Buy_i^*, \varepsilon_i) = 0$ .

To aid interpretation, note that if  $\rho = 0$  in (7) above, selection on unobservables is not a concern, because unobservable influences on the choice to buy gas have no bearing on the percentage difference in groceries purchased. At the other extreme (i.e. the upper bound),  $\rho$  takes on some positive value less than 1, and a positive shock towards being a gas buyer, results in correlated shock in the amounts of groceries purchased.

Having formalized the notation for the correlation term, we briefly discuss why the equality-of-effects assumption (i.e.  $\phi_{X'\gamma} = \phi_\varepsilon$ ) can provide an upper bound of the effect, in the presence of selection on unobservables (the formal arguments and proofs are presented in Altonji, Elder and Taber, 2002). Consider the universe of variables that completely explains the selection decision. Given our context, let us denote them as follows:

$$Gas\_Buy^* = X'_o\beta_o + X'_u\beta_u \quad (8)$$

Here, the subscript ‘o’ denotes “observed”, while the ‘u’ denotes “unobserved.” If the universe of covariates were partitioned such that  $X_o$  denoted a random partition, then the correlation of the index  $X'_o\beta_o$  with any outcome variable of interest should be approximately the same as the correlation of the index  $X'_u\beta_u$  with that same outcome variable. Since in reality, we do not observe  $X'_u\beta_u$ , these unobserved effects enter as a single quantity  $u$ , the residual term on the selection equation. Under random selection of observed covariates, we expect that the extent to which the index of the observables explains an outcome is about the same as the extent to which the unobservables explain the same outcome (equality-of-effects). Hence,  $\phi_{X'\gamma} = \phi_\varepsilon$ . This explains the intuition that underlies the upper bound of the correlation in Equation (7).

\*\*\* Table 4b \*\*\*

Table 4b presents the estimates of the AET model in (5). The upper bound of the correlation estimate ( $\rho = 0.093$ ) is very small, suggesting that the likelihood of unobservable selection swamping out the effect of gas station is small. Given this upper bound for the correlation, our most conservative estimate for the average percentage change in grocery spend in response to the opening of a gas station is 7.7%. We conclude that even accounting for selection, grocery spending among gas-buyers increases in the range of 7.7%-9.3% due to the gas station. The equivalent model for  $\Delta\%Trip_i$  (Column 2 of Table 4b) shows that the relative change in trips is close to estimates without accounting for unobservable selection (compare 15.3% versus 14.1% from

propensity-matched analysis in Table 4a). Overall, we conclude that our primary findings about the presence of demand externalities are robust to both observable and unobservable selection effects.

### 3.4 Profitability Implications of Expanding Scope

Should a retailer's decision to launch a gas station take into account the magnitude of the demand externalities on groceries? This would depend on relative magnitude of profits due to increased grocery sales, relative to the profits from gas. We perform a simple back of the envelope calculation to answer the question.

From the retailer, we know the amount of gasoline sold is approximately 30,000 gallons per week. At an average retail price of approximately \$2.85 per gallon on June 2005 to June 2006 in Connecticut (obtained from Gasbuddy.com), the revenue from gasoline is \$85,000 per week. At the upper bound of the externality (9.3% from propensity matching case in Table 4a), and with 39% of households buying gas, the unconditional increase in grocery spending is 3.6%. At the average weekly sales level of \$527,000 per week, this translates to additional revenues of \$19,000 per week. At the lower bound of the demand externality (7.7% after controlling both observable and unobservable selection in Table 4b), the unconditional increase in grocery spending is 3.0%, or \$15,700 per week. Thus the increased revenue from groceries is about 18-22% of revenues from gas.

Given that the percentage gross margin on groceries and gas are 35% and 5%, respectively (as indicated by managers at the grocery chain), the dollar gross margin is \$6,600 per week in increased groceries and \$4,300 per on gasoline sales at the upper bound of 9.3%. At the lower bound of 7.7%, the corresponding dollar gross margin from increased grocery sales is \$5,500 per week (gasoline margins remain the same). Thus, the increased profit from groceries is approximately 130% to 150% of the profits from gas.

On an annualized basis, the increased spillover profit from groceries ranges from \$286,000 to \$345,000, while the direct profits from gas is \$220,000. Interviews with managers indicated that the fixed costs of opening a gas station (during the period of the data) are approximately \$900,000. It is easy to see that the spillover benefits from groceries should have a significant impact on whether a retailer should expand scope by introducing a gas station.

Finally, we note that our analysis potentially understates the true spillover gains from increased scope in this context. We have only measured the extent to which *existing* customers change their behavior to drive revenue economies of scope. By design, such an analysis excludes two important additional effects: (1) the benefits for attracting new customers whose spending would be entirely incremental; and (2) the longer-term effects of improved customer retention. Both effects could further increase the demand externalities resulting from increased scope.

## 4. Robustness Checks

Our analysis thus far has estimated the change in behavior for households that choose to buy gas soon after the opening of the gas station. We now consider whether the results are robust to three sets of issues: First, are spillovers estimates positive merely due to a fortuitous timing of gas station entry, when demand tended to be high? Relatedly, could the positive spillover estimates be due to potential promotional activity following the gas station opening (which is unobservable to researcher)? Second, are the spillover estimates for early adopters of the gas station similar to the spillover for later adoption of the gas station? If the spillovers are due to the benefits of one-stop shopping, then the timing of adoption should not make a major difference between early and later adopters. Finally, are the estimated spillovers merely short-term effects or do they persist over the longer term?

To address the timing issue, we use the data over the entire 52 week period and estimate the following specification.

$$Groc_{it} = \alpha_i + \tau_t + \beta \cdot 1\{First\_Gas\_Use_i \leq t\} + \varepsilon_{it} \quad (9)$$

In this specification, rather than anchor the pre and post period around the opening of the gas station, we anchor the change at level of each household, where  $\beta$  captures the change in grocery spend after the household first uses the gas station. From the first use histogram in Figure 1, we see that the distribution of first use is quite spread out. The results indicate (Column 1 of Table 5) that the average increase in spending among gas-buyer households is \$5.44, similar to the estimates from our initial analyses, \$5.78, in Table 2. Our convergent results suggest that the spillover effects are not driven by the particular timing of entry; or due to particular promotional activities around the timing of the gas station.

\*\*\* Table 5 \*\*\*

We then extend the specification in Equation (9) to allow for different spillover effects depending on whether the household began using the gas station within the first 6 weeks or in later periods. Column 2 of Table 5 indicates there is no significant difference in spillovers between early and late adopters of the gas station.

Next, to test whether our results are not merely short-term effects, we test for differences in effects in six-week intervals after each household’s adoption of the gas station (Columns 3 of Table 5). Gas-buyers do appear to initially spend more on groceries relative to their incremental spend in subsequent weeks. However, there is still a positive and significant increase on grocery spending for the weeks remaining after an initial 18-week period following a household’s first use of the gas station, suggesting that longer-term effects do accompany the use of the new gas station.

Finally, we also assess the generalizability of the demand externality result, by testing whether the effect is replicated at the control store. We test whether the spend of “new” gas users who purchase after week 14 for the first time increase their grocery spend after they begin to buy gas. Table 6, Column 1 shows that indeed grocery spend is greater after first gas use. We next test whether these effects continue to persist over the longer term. From Table 6, column 2, we find that similar to the treatment store, there is a decline in subsequent weeks, but the spillover effects continue to persist over the longer term.

\*\*\* Table 6 \*\*\*

## 5. Moderators of Increased Grocery Spend

In this section, we seek a deeper understanding of the mechanism underlying the measured demand externalities by understanding moderators of the increased grocery spend. Specifically, we test (1) how the increased grocery spend is moderated by prior store loyalty and (2) how the percentage increase in spend varies across convenience store and grocery categories.

### 5.1 The Moderating Effect of Store Loyalty

Understanding how prior store loyalty moderates grocery spillovers from the gas station helps address the question of whether the gas station helps cement the loyalty of store loyal customers

or increase spending among casual shoppers. On the one hand, the more loyal the customer, the more likely such a customer might also patronize a new product offering; therefore the most loyal customers are those who are likely to respond to the new product offering and then increase their grocery spending through increased visits. On the other hand, those who are already loyal (i.e. high share of wallet) customers may simply not have further room to increase their grocery spending. In that case, it might be the moderate- or low-loyalty customers who respond most favorably to increased scope and increase their purchases on the firm’s original offering.

We model store loyalty through a “share of wallet” metric described in the data section. We estimate the following specification allowing for possible nonlinear effects of SOW on grocery spillovers:

$$\begin{aligned}
 \text{Groc}_{it} = & \alpha_i + \tau_t + \beta \cdot 1\{\text{Gas}_{Buy}\}_i \cdot 1\{\text{Gas}_{Open}\}_i \\
 & + \lambda_1 \text{SOW}_i \cdot 1\{\text{Gas}_{Buy}\}_i \cdot 1\{\text{Gas}_{Open}\}_i \\
 & + \lambda_2 \text{SOW}_i^2 \cdot 1\{\text{Gas}_{Buy}\}_i \cdot 1\{\text{Gas}_{Open}\}_i + \varepsilon_{it}
 \end{aligned} \tag{11}$$

As seen in Column 1 of Table 7, there is a concave relationship between SOW and the increase in grocery spending. This result is robust to the timing of households’ initial gas purchase (Column 2 of Table 7). Despite the concavity, the predicted absolute dollar increases in grocery spend per week are largest among the highest SOW households (Figure 2a).

\*\*\* Table 7 \*\*\*

Next, we describe the moderating effect of store loyalty in terms of percentage changes in grocery spend. For this, we normalize the grocery store spend by average weekly spend in the 14 weeks prior to opening the gas station. We report the percentage change in spend by SOW in Figure 2b. We find that the relative change in spending remains fairly constant above 50% SOW. Overall, we conclude that the gas station serves primarily to consolidate the loyalty of customers who are already loyal to the store.

## 5.2 The Moderating Effect of Categories

Understanding how grocery spillovers are moderated by convenience store categories or supermarket categories can give us insights about how the introduction of gas station impacts competitive market structure. First, does the increased grocery spending come primarily from

convenience stores, which are typically part of gas stations? If so, we would find increased spend only in convenience store categories. Second, does the one-stop shopping convenience due to the gas station lead to gain in grocery spend from other supermarkets? That is, does the gas station help strengthen the competitive position of the store against other supermarkets as well? If this is the case, we would also find increased spend in grocery categories as well. Overall, the analysis gives us insight into both intra-format and inter-format competition effects of the gas station.

Column 3 of Table 7, separate the incremental spend by grocery and convenience categories. Gas purchasers increase spending in both categories relative to non-gas buyers, with a 13% increase for grocery categories and 16% increase in the convenience categories.<sup>13</sup> Thus the incremental effect is about 23% larger for convenience categories than for traditional grocery categories. Thus the gas station improves the competitive position of the store both intra-format and inter-format. While the magnitude of the inter-format effect is larger given the direct competitive effect on gas stations on convenience stores, it is surprising that the magnitude of the intra-format effect (spend drawn from other supermarkets) is also roughly comparable. It suggests that at least in frequently purchased categories (such as gas) that are likely to be combined into one trip, the competitive benefits of providing one-stop shopping through the introduction of just one additional category can significantly change a supermarket's competitive position with respect to other supermarkets.

## 6. Conclusion

We demonstrate the existence of significant demand externality from co-location and measure the magnitude of this externality based on direct measures of household behavior in the context of a supermarket that opens a gas station in its parking lot. Our results are robust to controls for selection on observables and unobservables, and we estimate the short-term increase in grocery spend due to gas-station co-location to be between 7.7% and 9.3%. These increases in spending are accompanied by a 14%-15% increase in the number of trips made to the store, indicating that increased spending likely occurs due to the lower travel costs associated with one-stop shopping.

---

<sup>13</sup> Because of the systematic difference in dollar amounts between the two categories, we focus only on the percentage change in spending within the categories.

These spillovers are economically and managerially significant. Given our estimates, grocery spillovers are about 18-22% of gas revenues. Even more impressively, the gross profits from grocery spillovers are 130-150% of gross profits from gasoline sales.

The results are robust to concerns such as potential endogeneity about the timing of the gas station opening, and effect of unobservable events such as promotional activity around the time of the gas station opening. Moreover, the increase in spending on groceries is similar irrespective of when the household first uses the gas station and incremental spend continues to remain positive over the long run.

Finally, we find that households with higher loyalty levels show the greatest response to increases in scope, suggesting that the gas station helps to consolidate customer loyalty to the store. In terms of competitive impact, the gas station additional grocery dollars come at the expense of both traditional grocery competitors as well as convenience stores, though the inter-format effect on convenience stores is greater than the inter-format effects on other supermarkets.

The issue of demand externalities that arise from firm scope has been a major issue of research in economics and marketing. Our direct demand based household level approach to measure the externality complements traditional supply side approaches (based on supply decisions such as entry and prices), which require the assumption that firms behave optimally. We hope this research serves as an impetus for closer examination of demand externalities in a variety of contexts and informs broader the decision-making – ranging from decisions about retail scope and mall design to more ambitious goals such as the revival of downtown business districts.



## References

- Altonji, Joseph G, Todd E Elder and Christopher R Taber (2005), "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy*, 113(1), 151-184.
- Altonji, Joseph G, Todd E Elder and Christopher R Taber (2002), "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," Manuscript (April).
- Bailey, Tom Jr. (2009), "First Tennessee to Close 11 Grocery Store Banks by July", *Commercial Appeal*, May 7.
- Bell, David R. and Teck-Hua Ho (1998), "Determining Where to Shop: The Fixed and Variable Costs of Shopping," *Journal of Marketing Research*, 35(3), 352-369.
- Berger, Allen N., David B. Humphrey and Lawrence B. Pulley (1996), "Do Consumers Pay for One-Stop Banking? Evidence from an Alternative Revenue Function," *Journal of Banking & Finance*, 20, 1601-1621.
- Berger, Allen N., J. David Cummins, Mary A. Weiss and Hongmin Zi (2000), "Conglomeration vs. Strategic Focus: Evidence from the Insurance Industry," *Journal of Financial Intermediation*, 9, 323-362.
- Busse, Meghan, Jorge Silva-Risso, and Florian Zettlemeyer (2006), "\$1,000 Cash Back: The Pass-Through of Auto Manufacturer Promotions," *American Economic Review*, 96 (4), 1253-1270.
- Christofferson, Scott A., Robert S. McNish and Diane L. Sias (2004), "Where Mergers Go Wrong.," *McKinsey Quarterly*, (2), 92-99.
- Cummins, John David, Mary A. Weiss, Xiaoying Xie and Hongmin Zi (2010), "Economies of Scope in the Financial Services Industry: A DEA Efficiency Analysis of the US Insurance Industry" *Journal of Banking and Finance*, forthcoming.
- Datta, Sumon and K. Sudhir (2011), "The Differentiation-Agglomeration Trade-Off in Location Choice," Working paper, Yale University.
- DiPrete, Thomas A., and Markus Gangl (2004), "Assessing Bias in the Estimation of Causal effects: Rosenbaum Bounds on Matching Estimators and Instrumental Variables Estimation with Imperfect Instruments," *Sociological Methodology*, 24, 271-310.
- Frieswick, Kris (2005), "Fool's Gold" *CFO*, February.
- Gaffen, David (2007), "Five Reasons: Break Up Citigroup!" *The Wall Street Journal*, April 13, 2007.

- Gauri, Dinesh K., K. Sudhir and Debabrata Talukdar (2008), "The Temporal and Spatial Dimensions of Price Search: Insights from Matching Household Survey and Purchase Data," *Journal of Marketing Research*, 45 (April), 226-240.
- Gicheva, Dora, Justine Hastings and Sofia Villas-Boas (2010), "Revisiting the Income Effect: Gasoline Prices and Grocery Purchases," *American Economic Review: Papers and Proceedings*, 100 (2), 480-484.
- Goic, Marcel, Kinshuk Jerath and Kannan Srinivasan (2010), "Cross-Market Discounts," *Marketing Science*, forthcoming.
- Gould, Eric D, B Peter Pashigian and Canice J Prendergast (2005), "Contracts, Externalities, and Incentives in Shopping Malls," *The Review of Economics and Statistics*, 87(3), 411-422.
- Klemperer, Paul and A Jorge Padilla (1997), "Do Firms' Product Lines Include Too Many Varieties?" *The RAND Journal of Economics*, 28(3), 472-488.
- Leuven, Edwin and Barbara Sianesi (2003), "PSMATCH2: Stata Module to Perform Full Mahalanobis and Propensity Score Matching, Common Support Graphing and Covariate Imbalance Testing," <http://ideas.repec.org/c/boc/bocode/s432001.html>, version 4.0.4.
- Ma, Yu, Dinesh Gauri, Kusum Ailawadi and Dhruv Grewal (2011), "An Empirical Investigation of the Impact of Gasoline Prices on Grocery Shopping Behavior," *Journal of Marketing*, 75 (March), 18-35.
- Philips, Darsha (2010), "Target plans to open store in downtown L.A." *KABC-TV Los Angeles*, November 4, 2010.
- Rosenbaum, Paul R. (2002), *Observational Studies*, 2<sup>nd</sup> ed. New York, NY: Springer.
- Sun, Monic and Feng Zhu (2012), "Ad Revenue and Content Commercialization: Evidence from Blogs," Working Paper, University of Southern California.
- Vitorino, Maria Ana (2010), "Empirical Entry Games with Complementarities: An Application to the Shopping Center Industry," Working paper, The Wharton School.
- Wernerfelt, Birger (1994), "Selling Formats for Search Goods," *Marketing Science*, 13(3), 298-309.

**Table 1: Summary Statistics**

**(a) Summary Statistics – All Stores:** Weekly average spending and gas station use across all stores in data.

Summary Metric	Treatment Store (New Gas Station)	Control Store with Gas		
Number of Households	10,827	10,584		
Grocery Spend per Household per Week (52 weeks)	\$41.73	\$41.16		
Grocery Spend per Week Before Gas Station Opens (Weeks 1 - 14)	\$41.09	\$43.55		
Grocery Spend per Week After Gas Station Opens (Weeks 15-52)	\$41.96	\$40.28		
<b><i>Gas-Station Use</i></b>				
(First observed use of gas station)				
6 Weeks After Gas Station Opens	4,094	38%	5,468	52%
14 Weeks After Gas Station Opens	5,665	52%	5,992	57%
38 Weeks After Gas Station Opens (full data)	6,850	63%	6,811	64%

**(b) Summary Statistics – Treatment Store:** Weekly-average household grocery, gas and trip behavior, six weeks before and after first sale of gas.

Household Type	HH Count	% of Household	Grocery Sales				Gasoline Sales
			Pre-Period Spend	Post-Period Spend	Difference in Spend	%Change in Spend	Post-Period Spend
Gas Buyer	4,215	39%	\$50.78	\$54.36	\$3.58	7.1% ***	\$13.80
Non-Gas Buyer	6,675	61%	\$35.04	\$32.84	-\$2.20	-6.3% ***	\$0.00
Overall	10,890		\$41.13	\$41.17	\$0.04		\$5.34

Household Type	Grocery Trips		
	Pre-Period Trips	Post-Period Trips	Difference in Trips
Gas Buyer	1.07	1.13	0.06 ***
Non-Gas Buyer	0.73	0.66	-0.07 ***
Overall	0.86	0.85	-0.01

**Table 2: Change in Grocery Spending and Trips by Gas Buyers\***

Time Frame of Analysis	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	6 Weeks Before/After				14 Weeks Before/After		6 Weeks Before/After (Alternative Specifications)			
Dependent Variable	Diff. Grocery Spend	Pct. Diff. Groc.	Diff. Trips per Week	Pct. Diff. Trips	Pct. Diff. Groc.	Pct. Diff. Trips	Groc. Spend per Week	Groc. Trips per Week	Log [Spend per Week]	Groc. Trips per Week
Intercept	-2.20 *** (0.29)	-0.06 *** (0.01)	-0.06 *** (0.01)	-0.09 *** (0.01)	-0.02 ** (0.01)	-0.08 *** (0.01)	39.88 *** (0.43)	0.86 *** (0.01)	2.19 *** (0.02)	
Bought Gas at Treatment Store	5.78 *** (0.49)	0.13 *** (0.01)	0.13 *** (0.01)	0.15 *** (0.01)	0.15 *** (0.01)	0.15 *** (0.01)	5.66 *** (0.50)	0.12 *** (0.01)	0.28 *** (0.02)	0.14 *** (0.01)
<b>Fixed Effects</b>										
Week							Yes	Yes	Yes	Yes
Household							Yes	Yes	Yes	Yes
N	10,890	10,890	10,890	10,890	13,189	13,189	10,890	10,890	10,890	10,890
F-Value	136.52	134.10	209.62	234.28	236.28	327.87	32.27	56.08	46.90	
Prob>F	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	

\* p<.10, \*\* p<0.05, \*\*\* p<0.01

\* Difference regressions (columns 1, 2, 5, 6, 9 and 10) have robust standard errors and are weighted by household behavior in pre-period (gross grocery sales or trips). Panel regressions (columns 3, 4, 7 and 8) have clustered standard errors.

**Table 3: Percent Difference in Grocery Spending Among Gas-Buyers and Non-Gas Buyers**

<b>Time-Period of Analysis</b> (relative to opening of new gas station)		
<b>Pre-Period</b>	7-12 Weeks Before	1-6 Weeks Before
<b>Post-Period</b>	1-6 Weeks Before	1-6 Weeks After
<b>Treatment Store</b>	<b>All Households</b>	<b>Non-Gas Buyers</b>
<b>Control Store with Gas</b>	<b>None</b>	<b>Non-Gas Buyers</b>
Intercept	0.01 (0.01)	-0.09 *** (0.01)
Gas Buyer (Indicator)	0.01 (0.01)	
Store with New Gas Station (Indicator)		0.02 * (0.01)
N	10,404	12,131
F-Value	0.8522	1.0188
Prob>F	0.356	0.3128

\*  $p < .10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

\* Percentage difference in grocery spend regressions are weighted by total spending in the pre-period. Robust standard errors in parentheses.

**Table 4a: Robustness to Selection on Observables - Propensity-Matched Results:<sup>\*</sup>  
Increase in Weekly Grocery Spending & Trips**

Propensity-Matching Variable(s)	Diff. Groc (Post - Pre)	Pct. Diff. Groc (Weighted)	Diff. Trips (Post - Pre)	Pct. Diff. Trips (Weighted)
(1) Results w/o matching (Tables 2 & 3)	5.778 (0.495)	0.133 (0.012)	0.127 (0.009)	0.147 (0.010)
(2) Spend, Trips	7.038 (0.514)	0.096 (0.010)	0.160 (0.009)	0.146 (0.008)
(3) log(Potential Spend), Trips, Pct. SoW	6.850 (0.515)	0.093 (0.010)	0.158 (0.009)	0.141 (0.008)

<sup>\*</sup> All results are significant with p-value of <0.01. “Spend” indicates grocery spending in the pre-period. “Potential Spend” is grocery spend in the pre-period divided by the percentage share of wallet (“Pct. SoW”, above). Standard errors are in parentheses. Percentage changes are weighted by household grocery spending or trips in the pre-period (as relevant).

First row of results are identical to those reported in Tables 2 and 3 (coefficient for “Bought Gas”), and reflect estimated change in behavior without any propensity matching.

Propensity scores reflect the probability of a household buying gas, and are estimated using a linear probit model. Matching is implemented using Gaussian kernels (bandwidth of 0.05) on variables as shown. Numbers in the table above should be understood as the average treatment effect *on the treated*.

**Table 4b: Robustness to Unobserved Selection Effects:<sup>\*</sup>**  
**Percent Difference in Grocery Spending**

Variable	Percent Diff. Groceries	Percent Diff. Trips
<b>Outcome Equation</b>		
Gas_Buy	0.077 * (0.033)	0.153 *** (0.028)
Intercept	1.580 *** (0.171)	0.876 *** (0.029)
log(Potential)	-0.247 *** (0.024)	-0.127 *** (0.004)
Trips	0.0026 (0.004)	-0.0058 *** (0.002)
SoW	-0.0378 (0.072)	-0.2314 *** (0.011)
Sigma	0.6162 ** (0.230)	0.5445 *** (0.005)
<b>Selection Equation (Gas_Buy*)</b>		
Intercept	-1.1739 *** (0.313)	-0.4475 *** (0.072)
log(Potential)	0.0944 * (0.042)	-0.0140 (0.013)
Trips	0.0521 *** (0.009)	0.0671 *** (0.004)
SoW	0.2463 (0.134)	-0.2442 *** (0.033)
<b>Correlation (restricted)</b>	0.0933	0.0550

<sup>\*</sup> Maximum-Likelihood estimates with restrictions on correlation term. Observations are weighted by grocery spend or trips in the pre-period (as relevant). Standard errors are obtained from 200 bootstrap iterations.

**Table 5: Longer Term Changes in Grocery Spending by Gas Buyers \***

Dependent Variable	(1)	(2)	(3)
	Spend per Week	Spend per Week	Spend per Week
Intercept	40.38 *** (0.45)	40.38 *** (0.45)	40.38 *** (0.45)
Gas User (Individual Timing)	5.44 *** (0.29)	5.55 *** (0.37)	6.53 *** (0.33)
Incr. Grocery Spend if First Gas Use in Weeks 7-12 After New Gas Opens	- -	-0.15 (0.60)	- -
Incr. Grocery Spend if First Gas Use in Weeks 13-18 After New Gas Opens	- -	-0.66 (0.92)	- -
Incr. Grocery Spend if First Gas Use in Weeks 19+ After New Gas Opens	- -	0.90 (1.04)	- -
Change in Incr. Spend for New Users 7-12 Weeks After First Gas Use	- -	- -	-1.12 ** (0.37)
Change in Incr. Spend for New Users 13-18 Weeks After First Gas Use	- -	- -	-0.92 * (0.36)
Change in Incr. Spend for New Users 19+ Weeks After First Gas Use	- -	- -	0.00 (0.35)
<hr/>			
<b>Stores</b>	Treatment Store	Treatment Store	Treatment Store
<b>Fixed Effects</b>			
Week	Yes	Yes	Yes
Household	Yes	Yes	Yes
F-Value	61.01	57.74	59.46
Prob>F	0.0000	0.0000	0.0000

\* Clustered standard errors in all regressions.



**Table 6: Changes in Grocery Spending at the Control Store with Gas**

Dependent Variable	(1)	(2)
	Spend per Week	Spend per Week
Intercept	40.78 *** (0.55)	40.80 *** (0.55)
Incumbent Gas Users - (i.e. First Gas Purchase On/Before Week 14)	3.44 *** (0.59)	3.40 *** (0.59)
"New" Gas Users - (i.e. First Gas Purchase On/After Week 15)	4.05 *** (0.53)	6.98 *** (0.66)
Change in Incr. Spend for "New" Users 7-12 Weeks After First Gas Use	- -	-3.39 *** (0.78)
Change in Incr. Spend for "New" Users 13-18 Weeks After First Gas Use	- -	-0.76 (0.81)
Change in Incr. Spend for "New" Users 19+ Weeks After First Gas Use	- -	-1.06 (0.79)
<b>Fixed Effects</b>		
Week	Yes	Yes
Household	Yes	Yes
N	10,584	10,584
F-Value	69.52	66.45
Prob>F	0.0000	0.0000

\* p<.10, \*\* p<0.05, \*\*\* p<0.01

**Table 7: Moderators of Increased Spending by Gas Buyers at Treatment Store**

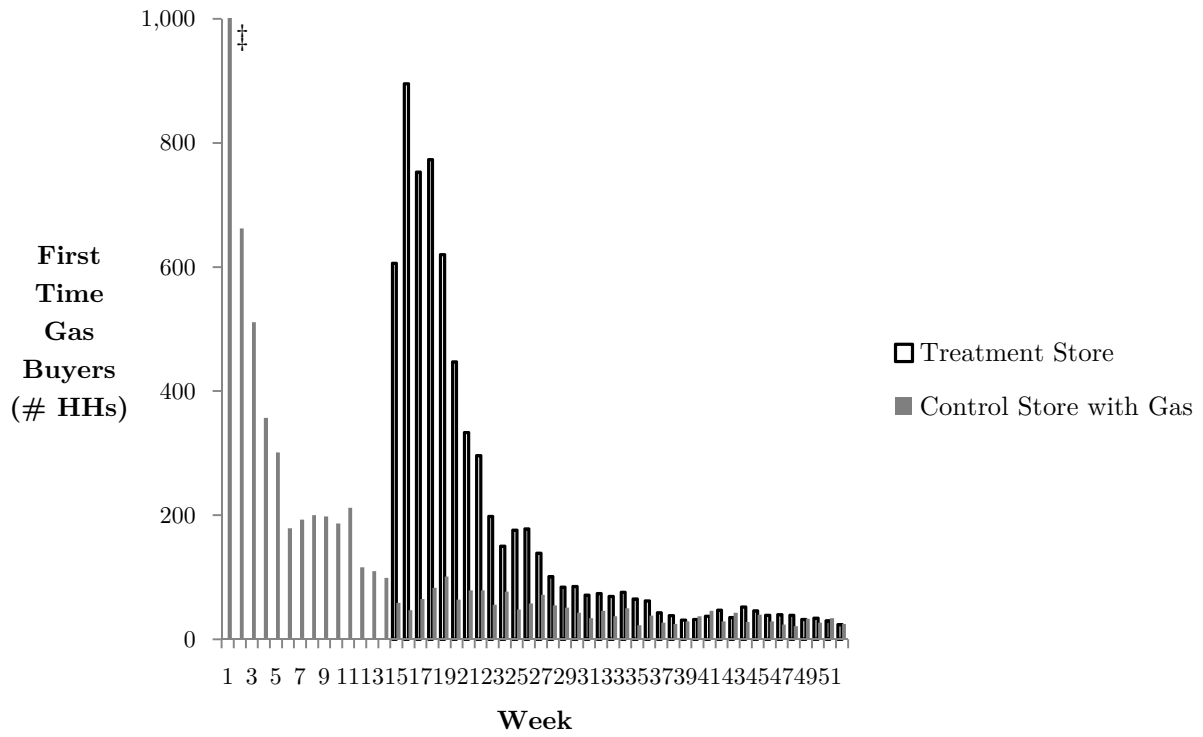
Dependent Variable	(1)	(2)	(3)	(4)	(5)
	Spend per Week	Spend per Week	Pct. Diff. Spend Traditional	Pct. Diff. Spend Convenience	Pct. Diff. Spend All Groceries
Intercept	40.38 *** (0.45)	40.38 *** (0.45)	-0.07 *** (0.01)	-0.05 *** (0.01)	-0.07 *** (0.01)
Gas User	-3.35 *** (0.58)	-2.71 *** (0.53)	0.13 *** (0.01)	0.15 *** (0.01)	0.13 *** (0.01)
Gas User × Pct. SOW	24.48 *** (2.88)	21.93 *** (2.82)	-	-	-
Gas User × Pct. SOW Squared	-11.00 *** (2.80)	-8.06 ** (2.75)	-	-	-
Convenience Products	-	-	-	-	0.01 (0.01)
Gas User × Convenience Products	-	-	-	-	0.03 * (0.01)
<b>Gas User Status</b>	Fixed	Variable			
<b>Fixed Effects</b>			N/A	N/A	N/A
Week	Yes	Yes	6 Weeks Pre/Post	6 Weeks Pre/Post	6 Weeks Pre/Post
Household	Yes	Yes			
N	10,827	10,827	10,865	10,580	21,445
F-Value	61.08	63.88	115.49	108.26	52.73
Prob>F	0.00	0.00	0.00	0.00	0.00

\* p<.10, \*\* p<0.05, \*\*\* p<0.01

\* Panel regressions (columns 1 and 2) have clustered standard errors. Gas User Status indicates how a household is understood to be a gas user. In the Fixed case (column 1), a household who purchased gas is denoted as a gas user for all observations after the gas station opens, regardless of the timing of that household's first gas purchase. In the Variable case (column 2), a household only becomes a gas user for data points on and after the first week of observed gas purchase.

Percentage difference regressions (columns 3 through 5) have robust standard errors and are weighted by household behavior in pre-period (gross grocery sales). Column 3 indicates percentage change in spending for traditional grocery items (Trad.), while column 4 indicates the same change for convenience items (Conv.). Column 5 includes data for both traditional and convenience items (All).

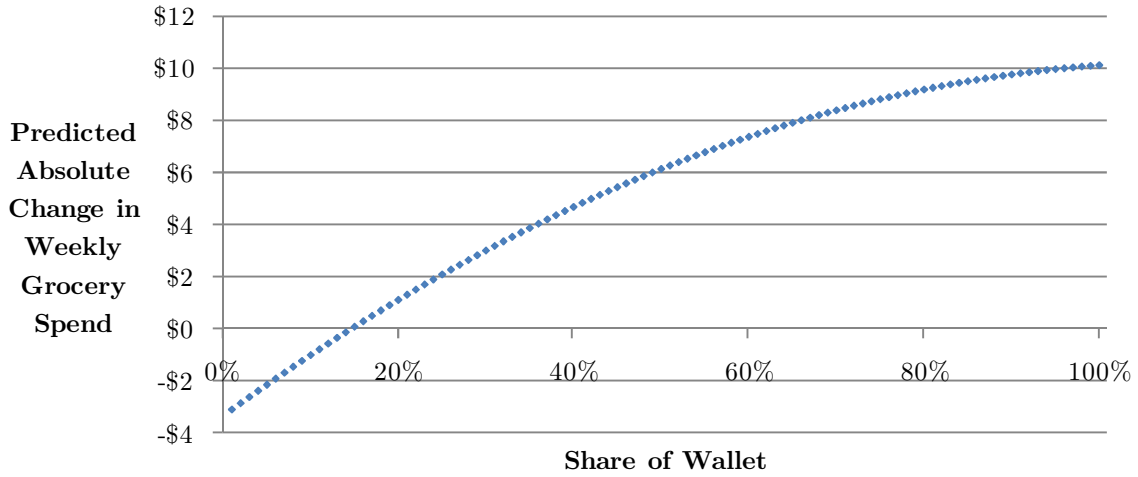
Figure 1: Histogram of First Gas Station Use by Week



‡ Household count for first week in store with pre-existing gas station is 1,724.

**Figure 2: Moderating Effect of Household Loyalty on Grocery Spending**

**(a) Predicted Change in Grocery Spend for Gas Users, by Share of Wallet:** Predicted results plotted based on estimates in Column 2 of Table 6.



**(b) Calculated Relative Change in Grocery Spend for Gas Users, by Share of Wallet:** Plotted values represent predicted change (from above figure) divided by average weekly spend in 14 weeks before gas station opens for households that eventually use the gas station.

