A Mixed Bag: The Hidden Time Costs of Regulating Consumer Behavior

Rebecca Taylor, University of Sydney, r.taylor@sydney.edu.au

Selected Paper prepared for presentation at the 2018 Agricultural & Applied Economics Association Annual Meeting, Washington, D.C., August 5-August 7

A Mixed Bag:

The Hidden Time Costs of Regulating Consumer Behavior*

Rebecca Taylor†

April 28, 2018

Abstract

The non-monetary costs consumers experience from price and quantity regulations are challenging to quantify, and thus easily overlooked. Using quasi-experimental policy variation and high-frequency supermarket data, this paper identifies time costs of policies that ban and tax the use of disposable carryout bags (DCB). DCB policies cause persistent 3% increases in supermarket checkout duration. Moreover, DCB policies and their associated time costs disincentivize grocery shopping, with supermarkets in regulated jurisdictions experiencing a 1.6% reduction in sales. The results explicitly link time costs as a mechanism behind policy avoidance and highlight trade-offs between convenient and environmentally-friendly behaviors in healthy food acquisition.

JEL Codes: D12, D62, H23, Q52, Q58

*I am indebted to Sofia Berto Villas-Boas, Peter Berck, Hilary Hoynes, and David Zilberman for their advice and mentorship. I thank Michael Anderson, Maximilian Auffhammer, Kendon Bell, Joshua Blonz, Fiona Burlig, Aluma Dembo, Meredith Fowlie, Stephanie Heger, Leslie Martin, Jeffrey Perloff, Louis Preonas, John Schindler, Andrew Stevens, Daniel Tregeagle, Erin Wolcott, Yang Xie, as well as the participants of the UC Berkeley ERE seminar and the WEAI Dissertation Workshop, for helpful discussions and suggestions. I thank Kate Adolph, Katherine Cai, Samantha Derrick, Tess Dunlap, Valentina Fung, Claire Kelly, Ben Miroglio, Nikhil Rao, Lucas Segil, Corinna Su, Edwin Tanudjaja, and Sarah Zou for their superb research assistance. This project would not be possible without the institutional and technical support of the retailers that provided data and access to their stores. This paper reflects the author's own analyses and calculations based on data from individual retailers and data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data at the University of Chicago Booth School of Business. The conclusions drawn from the Nielsen data are those of the researchers and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

†Mailing address: Room 370, Merewether Building [H04], The University of Sydney, NSW 2006, Australia. Tel: 1 513 600 5777. Email: r.taylor@sydney.edu.au.

1 Introduction

Governments often enact policies to incentivize consumers away from behaviors with negative externalities, at the expense of consumer convenience. Increases in tobacco, alcohol, and sugar taxes incentivize consumers to spend time and effort reoptimizing their consumption baskets away from these products (Becker et al., 1991; Chaloupka, 1991). Seat belt laws and speed limits incentivize drivers to spend time and effort buckling up and slowing down. Energy efficiency subsidies (Fowlie et al., 2015; Allcott, 2016), garbage pricing (Fullerton and Kinnaman, 1996), and bottle return refunds (Beatty et al., 2007; Ashenmiller, 2011) incentivize consumers to spend time and effort in conserving energy, reducing waste, and recycling. These policies illustrate that when consumer convenience is at odds with public health, safety, and the environment, policymakers "ask" consumers to trade personal convenience to benefit the public good. This begs the empirical questions: how large and persistent are the time, effort, and psychological costs consumers trade in changing their behavior, and how do consumers respond to these non-monetary costs?

Building on Becker's (1965) theory of the allocation of time, the economic literature widely acknowledges the importance of time costs for conducting welfare analyses[1] and improving policy design.[2] However, in many settings quantifying these costs is infeasible, causing them to be overlooked. This is particularly true when the behaviors to be altered are frequent but individually short-lived. In this paper, I explore a hidden time cost of policies aimed at altering consumer behavior. Specifically, I examine how local government

[1]For instance, Krueger (2009) estimates that Americans spent more than 800 million hours waiting to obtain medical services in 2007. The omission of patients' time in expenditure calculations caused national health care expenditures to be undercounted by 11 percent.

[2]Numerous studies have shown how defaults can be set recognizing the time costs of opting out, such as in the case of retirement savings (Madrian and Shea, 2001; Choi et al., 2003) and organ donation (Johnson and Goldstein, 2003). In the technology adoption literature, a frequent conclusion is that consumers may not adopt a privately beneficial product, even when it is free, if the non-monetary costs of obtaining or using the product are high (Dupas, 2014; Fowlie et al., 2015).

regulation of disposable carryout bags (DCB) affects the wait and processing time of checkout services provided by supermarkets. While DCBs bring convenience to supermarkets and supermarket customers, they are costly to the environment and to governments trying to keep their streets and waterways clean. In order to curb the consumption of single-use bags and encourage the use of reusable bags, DCB policies prohibit retail stores from providing customers with "free" bags at checkout.[3] Using high-frequency supermarket scanner data and variation in DCB policy adoption over time and space in an event study design, this paper addresses two fundamental questions: 1) What are the time costs to consumers of policy-induced behavioral change, and 2) How do time costs alter future consumption decisions?

Several features of supermarket checkout and DCB policies make them a compelling setting to study the time costs of changing behavior. First, food shopping is a common, frequent, and arguably necessary behavior. In the U.S., consumers purchase the majority of food at grocery stores, supermarkets, and superstores (Taylor and Villas-Boas, 2016a), with the average adult grocery shopping once every 7.2 days and spending 44 minutes in-store per trip (Hamrick et al., 2011). This aggregates to 11.9 billion grocery shopping trips (and checkout experiences) and 8.7 billion hours in-store each year in the U.S.[4] Given the extensive literature showing that shopping convenience impacts where and what people purchase to eat,[5] and the literature showing that consumers dislike and actively avoid long wait times (Larson, 1987; Tom and Lucey, 1995; Van Riel et al., 2012; Craig et al., 2016), it is vital that we understand the trade-offs between convenient behaviors and environmentally-friendly behaviors in healthy food acquisition.[6]

[3]Supermarkets pass the cost of disposable carryout bags on to their customers in the overall price of groceries.

[4]Author's calculation using population data from the 2010 United States Census.

[5]See Yaktine and Caswell (2013) for a comprehensive review of this literature.

[6]People place a higher value on time spent waiting than they do on the same amount of time in other

Second, supermarket checkout is a setting where changes in the time spent in an activity can be precisely identified and measured. With time-stamped, transaction-level scanner data obtained from a national supermarket chain, I know exactly when and where a checkout transaction occurred (e.g., register 2 in store X and city Y on Saturday, April 27, 2013 at 2:07pm), who was present (e.g., cashier A and customer B), what was purchased (e.g., two boxes of cereal C at \$3.79 each), and importantly, how much time the transaction took to complete. Unlike many time-use studies, these data do not rely on surveys and time diaries, which can be expensive to implement and are prone to systematic under/over reporting and recall bias (Neter and Waksberg, 1964; Mathiowetz and Duncan, 1988). Moreover, the panel nature of the scanner data allows me to examine the effects of the policy change over time, at the store, cashier, and customer level.

Third, DCB policies are widely used legislative tools for changing consumer behavior. With growing concern over the millions of dollars spent per year in DCB clean-up, recycling, and landfilling,[7] lawmakers across the country have adopted DCB policies to change how their constituents obtain food. From 2007 to 2016, approximately 242 local governments adopted DCB policies across 20 states and the District of Columbia.[8] While lawmakers acknowledge the trade-off between convenience and the environment in regulating prices and quantities,[9] DCB policies typically have been evaluated based on the magnitude of behavioral

circumstances (Maister, 1985; Small and Verhoef, 2007). Not only do long lines have a time cost, they also have an emotional one: "stress, boredom, that nagging sensation that one's life is slipping away."("Why Waiting is Torture." *New York Times*. Aug. 19, 2012. Online, *accessed Mar. 25, 2016.*)

[7]Local governments are estimated to spend 1.1 cents per bag in collection, processing, and landfilling costs (Herrera Environmental Consultants, Inc., 2008). Given that approximately 100 billion plastic bags are consumed in the U.S. each year (Clapp and Swanston, 2009), municipalities nationwide spend \$1.1 billion per year to manage plastic bags. This taxpayer estimate does not include the environmental cost of plastic once it enters oceans.

[8] For lists of disposable bag policies in the U.S.—by city, county, and state (and by adoption/rejection date)—see: BagLaw.com and Californians Against Waste, *accessed Sep. 5, 2016.*

[9]For instance, the City of Portland states, "Single-use plastic carryout bags may offer short-term convenience, but they have long-term costs." (Online, *accessed Sep. 10, 2016*).

change and litter reduction, and not full social welfare. Thus, even though several studies have found bans and taxes to be quite effective in altering consumer bag choices,[10] little is known about how these policy-induced behavioral changes affect the time and effort of individual consumers.

To understand the mechanisms behind the time cost of DCB policies, it helps to think about the production function of supermarket checkout. DCB policies may increase the duration of checkout through each of the three main inputs into the production function—namely, bags, labor, and capital. With respect to bags, DCB policies directly change the choice set of bags and their prices, and different bags vary in packing time. With respect to labor, DCB policies force cashiers to learn new key codes and procedures for collecting fees. Cashiers and baggers must also ascertain the number and types of bags customers want, and how to pack them. If customers do not bring bags, customers must decide how many bags for which they are willing to pay. This turns a decision that was automatic and habitual (i.e., fast thinking) into an economic, utility maximizing decision (i.e., slow thinking) (Kahneman, 2011). Finally, with respect to capital, checkout lanes are optimized for single-use plastic bags, with plastic bag dispensers placed next to the cashier and at the end of the bagging area. Importantly, checkout machinery is fixed in the short-run, and retailers cannot instantaneously reoptimize their machinery for alternative bags. Therefore, there exist several mechanisms through which DCB policies could lead to longer checkout wait and processing time, some of which may be reduced over time through learning-by-doing and learning-by-using (Arrow, 1962; Rosenberg, 1982). My analyses will shed light into each

[10] DCB policies come in two flavors: bag bans—command-and-control approaches to regulate behavior directly—and bag fees/taxes—market-based approaches to incentivize individuals to change their own behavior. Taylor and Villas-Boas (2016b) find that a plastic bag ban in California led to a 26 percentage point (ppt) increase in the use of reusable bags and a 9ppt increase in the use of no bags. However, the eradication of plastic bags was offset by a 47ppt increase in paper bag use. Homonoff (2016) studies the impact of a plastic and paper bag tax in Maryland and finds that the share of transactions using disposable plastic bags declined by 42ppt after the tax implementation. Additional studies have found DCB policies to be effective in changing bag choice in Ireland (Convery et al., 2007) and South Africa (Dikgang et al., 2012).

of these mechanisms.

To identify the time cost of DCB policies on checkout duration, I exploit a quasi-experiment in California, where city and county adoption of plastic bag bans has varied across both time and space. Leveraging this spatial and temporal variation to control for potentially confounding factors, I employ an event study empirical strategy. The event study model identifies the time cost of bag bans on checkout duration by comparing checkout duration at stores in jurisdictions with policies to checkout duration at stores in jurisdictions yet to be treated and in jurisdictions that are not treated during the sample. Importantly, plotting the differences between treated and control supermarkets over event-time enables me to directly test the identifying assumption of parallel trends in the pre-policy period, and to explore the dynamics of the policy effects in the post-policy period.

The event study results reveal that plastic bag bans cause a 3% average increase in checkout duration. I document heterogeneity in the policy effects by whether a customer chooses to pay for paper bags in lieu of plastic bags, with transactions paying for paper bags experiencing a 9.2% slowdown and transactions not paying for paper bags experiencing a 1.7% slowdown. Surprisingly, even though I observe evidence of learning at the cashier level, this learning does not eliminate the slowdown from the policies, which persists over the entire sample period. While 3% slower checkout durations (or roughly 3.6 seconds more per customer) may seem negligible, this time cost aggregates to 1.4 million hours across all grocery shopping trips made per year in California.

In a second set of event study results, using additional data on retail sales per store and month, I show that plastic bag bans and their corresponding time costs disincentivize grocery shopping, with supermarkets and grocery stores in treated jurisdictions experiencing a persistent 1.6% drop in sales. If the stores in the retail sample are representative of all supermarkets and grocery stores in California, a statewide DCB policy would cause a

\$2.3 billion loss in grocery sales annually. As a comparison, Californians received roughly \$6.7 billion in federal Supplemental Nutrition Assistance Program benefits (formerly Food Stamps) each year.[11] I then explicitly test the link between checkout time costs and lost sales. A hazard model analysis of how transaction duration affects the likelihood that consumers return to a store in subsequent weeks shows that 10% of the reduction in supermarket sales from DCB policies can be attributed to the 3% slowdown in transaction duration. Thus not only do DCB policies impose a time cost on customers, customers respond to this time cost by shopping less frequently at affected stores. Assuming people are still eating the same amount of food as before, where are these lost grocery sales going? I provide suggestive evidence that sales at restaurants and other food-away-from-home establishments increase when DCB policies go into effect. This unintended consequence is concerning because away-from-home meals tend to be higher in energy-density, fat, and sodium and lower in fruits, vegetables, and whole grains than at-home meals (Kant et al., 2015).

I conduct a series of robustness checks to further explore the results and their external validity. First, I test the robustness of the scanner data results to the use of an alternative data source—observational data collected in-store before and after a DCB policy change. I estimate results consistent with the scanner data, demonstrating that missing variables in the scanner data (i.e., the presence of baggers, the types of bags purchased, and the gender and race of the customer) are not biasing my results. Second, I replicate the analysis on data from an alternative store chain—a regional discount chain targeting low-income and bargain shoppers. I show the effects of DCB policies are not unique to the main retail chain in this paper. Third, I replicate the analysis using the 2010 Washington DC bag tax—a policy in a different location, with a different regulation tool. With scanner data from stores

[11] Source: "The CalFresh Food Assistance Program." *Public Policy Institute of California*. Feb. 2018. Online, *accessed Apr. 26, 2018.*

in the DC metropolitan area, I again observe persistent checkout slowdowns due to the policy change. However, unlike the California bag bans, the slowdown from the DC bag tax lessens significantly over time, suggesting bag choice is an important mechanism behind the slowdown.

This paper contributes to an emerging literature on the hidden costs of changing consumer behavior. Allcott and Kessler (2018) evaluate the welfare effects of social comparisons in reducing energy consumption and show that ignoring the non-energy costs of the intervention overstates its welfare gains by a factor of five. Similarly, Damgaard and Gravert (2016) study the annoyance cost of a nudge intervention and show that when not accounting for the non-monetary costs of reminders, the average welfare effects are overstated by a factor of ten. However, a limitation of these studies is that the non-monetary costs of the nudges are not directly measured, but instead are imputed using consumers' avoidance of the intervention (or willingness-to-pay to avoid the intervention). There is also a rich literature estimating how consumers circumvent regulations in unintended ways,[12] yet this literature often stops short of identifying the specific mechanisms behind the circumvention. Thus, this study is the first to simultaneously (i) quantify the time cost of a policy change directly and separately from other non-monetary costs, (ii) causally identify consumer avoidance of the policy, and (iii) explicitly link the observed time cost as a mechanism behind the unintended avoidance behavior.

This paper highlights the tension between convenient behaviors and environmentally-friendly behaviors in healthy food acquisition. Economic incentives and regulations which seem like low-cost behavioral nudges, may have large non-monetary costs with respect to time and convenience when aggregated across all consumers and all consumption occasions.

[12] For example, an unintended consequence of smoking bans in public places is that they displace smokers to private places where they contaminate non-smokers, especially young children (Adda and Cornaglia, 2010). An unintended consequence of driving restrictions by license plate number is that consumers circumvent the policy by increasing the total number of vehicles in circulation (Davis, 2008).

Moreover, not understanding the non-monetary costs of a policy can lead to unintended and perverse consequences if consumers try to avoid inconvenient policies with riskier or more harmful behavior. While often challenging to measure, and thus easy to overlook, quantifying these costs is vital for accurate welfare analysis and improved policy design.

2 Setting, Research Design, and Data

2.1 Background on Disposable Carryout Bags and Regulations

When first invented, plastic carryout bags were considered an engineering feat: "a waterproof, durable, featherweight packet capable of holding more than a thousand times its weight" (Freinkel, 2011). However, the characteristics that make plastic bags convenient also make them costly to the environment and to municipalities trying to keep their streets and waterways clean. Their lightweight and aerodynamics make it easy for them to blow out of waste streams and into the environment, where, due to their durability and water-resistance, they persist for a long time.

Each year Americans consume approximately 100 billion plastic carryout bags (Clapp and Swanston, 2009)—over 300 bags per person per year. Local governments are estimated to spend 1.1 cents per bag in clean-up, processing, and landfilling (Herrera Environmental Consultants, Inc., 2008), which aggregates to municipalities nationwide spending $1.1 billion per year.[13] In addition to clean-up costs, there are environmental costs of plastic bags if they escape their local jurisdictions. Jambeck et al. (2015) calculate that 1.7-4.6% of the plastic waste generated in coastal countries around the globe is mismanaged and enters the ocean. Once in waterways, plastic bags do not biodegrade, but instead break into smaller pieces,

[13]The CA Senate Rules Committee (CSRC, 2014) cite a lower estimate, with Californian taxpayers spending $25 million per year to dispose of 14 billion plastic bags, which is $0.002 per bag. Conversely, a study of the budgets of six major cities in the U.S. cites a higher estimate, with litter control costs of $0.032 and $0.079 per bag (Burnett, 2013).

which can be consumed by fish, turtles, and whales that mistake them for food.[14] Given both the environmental and clean-up costs of DCBs, lawmakers across the country are turning to policies to regulate their use. As of December 2016, approximately 242 local government DCB policies had been adopted across 20 states and the District of Columbia.[8] DCB policies employ one or both of the following policy tools to alter consumer behavior: (1) bag bans (i.e., quantity regulations), and (2) bag fees (i.e., price regulations).

California provides a rare quasi-experiment for analyzing the effects of DCB policies on checkout duration and grocery sales. In California, DCB policies ban retail food stores from providing customers with plastic carryout bags under 2.25 mils thick (i.e., traditional plastic carryout bags) and require stores to charge a minimum fee for all paper and reusable carryout bags provided at checkout.[15,16] From 2007 through 2014, DCB policies were implemented in 111 Californian cities and counties, affecting one third of California's population.[17] This local legislative momentum continued and culminated with the nation's first statewide plastic carryout bag ban, which was voted into law on November 8, 2016.[18]

Figure 1 maps the implementation of DCB policies in California at four points in time.

---

[14] A survey of experts, representing 19 fields of study, rank plastic bags and plastic utensils as the fourth severest threat to birds and marine animals in terms of entanglement, ingestions, and contamination (Wilcox et al., 2016). While plastic bags and films represent only 2.2% of the total waste stream (CSRC, 2014), plastic grocery bags and other plastic bags are the eighth and sixth most common item found in coastal cleanups ("International Coastal Cleanup. Annual Report 2016." *Ocean Conservancy.* Online, *accessed Aug. 31, 2017*).

[15] While the majority of DCB policies in California require a 10-cent fee for paper and reusable bags, a handful of jurisdictions have opted for either no fee, a 5-cent fee, or a 25-cent fee.

[16] Why has California chosen bag bans over bag fees? CA Assembly Bill 2449, enacted in 2006, began as a plastic bag fee bill, but due to pressure from the plastic industry, transformed into a plastic bag recycling bill. Additionally, this bill temporarily prohibited any public agency from adopting a regulation that imposed a plastic bag fee upon a store. Consequently, a bag fee was not an available policy tool for local governments in California that wanted to regulate plastic bags. ("The Plastic Bag Ban Epic." *LA Observed.* Sep. 6, 2014. Online, *accessed Oct. 9, 2016*).

[17] Author's calculations. See Online Appendix Table A.1 for a list of California DCB policies and implementation dates from 2007 to 2014.

[18] Californians voted and passed the Plastic Bag Ban Referendum (Proposition 67) by a margin of 52.9% (yes) to 47.1% (no).

Similar to other local government waste regulations, DCB policies may be implemented by city councils (for incorporated areas), county boards of supervisors (for unincorporated areas), and county waste management authorities (for entire counties with opt-out options for incorporated areas). City-level policies are depicted with dark green circles. Unincorporated county policies are shaded in light yellow. Countywide policies—where all unincorporated areas and all cities in a county implement DCB regulations—are shaded in dark green. This figure shows that DCB policies have varied greatly across both implementation dates and locations.

It is important to note that local jurisdictions decide when DCB policies will be operative; the stores within a jurisdiction do not make this decision. The start date is specified in a jurisdiction's ordinance document (i.e., bill). Examining the ordinance documents of all 111 jurisdictions in California that implemented DCB policies between 2007 and 2014, I find that 21% of jurisdictions specified January 1 as their start date, 30% specified the first of a month that was not January, 14% chose Earth Day (April 22), 11% chose a specific date other than the first of the month, and 23% did not specify a specific date and instead wrote to be operative 1, 3, or 6 months after adoption. Start dates vary across all days of the week. Importantly, while start dates were not randomly chosen, the dates were also not selected in a systematic way across all jurisdictions. My event study empirical strategy exploits this quasi-random variation in DCB policies across time and space to explore how DCB policies influence checkout duration and supermarket sales.

2.2 Sample Selection for Scanner Data

Quantifying the shock to checkout duration from DCB policies requires a detailed dataset on the speed and location of checkout transactions. To this end, I obtained access to time-stamped scanner data from a large supermarket chain.[19] While this retail chain processes as many as 800,000 items per hour in California alone, there was a limit to the amount of data I could request at the transaction level. Thus I designed a subset of scanner data, selecting data from comparable treated and control stores across California between January 2011 and May 2014.

My procedure for selecting the stores was as follows. First, the retailer provided a list of their stores in California with basic characteristics, such as street address, city, zip code, date opened, last date remodeled, and building area size. I merged in store level demographic data, created by Gicheva et al. (2010) using 2000 US Census data for each store's census block-group. Next, I split the sample of stores into treated stores and control stores, using the database of DCB policies I constructed for California.[17] Control stores are defined as stores in jurisdictions (i.e., incorporated cities and unincorporated counties) that have not implemented a policy during the sample period, whereas treated stores are stores in jurisdictions that have implemented a policy during the sample period. As a first step in ensuring that control stores are good counterfactuals for treated stores, I dropped stores in counties where no DCB policies had been adopted and stores in counties more than 50 miles from the coast.[20] As a second step, I used a propensity score matching algorithm, based on store age, store size, and stores' census block-group characteristics, to select 30 pairs of treated and control stores with sufficient overlap in observables. The data request for these stores was submitted in July 2013. By the time the request was approved and the data were pulled in May 2014, additional DCB policies had been enacted, causing 8 of the original control stores to be categorized as treated. Furthermore, after receiving the data I decided

[19] There are over 2000 locations of this supermarket chain across the U.S. With revenue over $35 billion per year, this chain is one of the 15 largest retailers in the U.S.

[20] Thus control stores are in counties where other jurisdictions have implemented policies, but their own jurisdiction has not.

to drop 7 stores which experienced either closure, remodeling, or policy differences, as these events could confound my checkout productivity measures.[21] Thus, in the final sample I have 53 stores—33 treated and 20 control—across 45 policy jurisdictions.

Importantly, the treated stores were chosen to mirror the variation of policy implementation dates in California. Figure 2 presents the number of municipalities in California implementing a DCB policies (depicted by the gray bars) and the number of stores in my sample in jurisdictions implementing a DCB policy (depicted by the black bars) in each month over the sample period—which spans January 2011 through May 2014. As designed, the distribution of implementation dates for stores in my sample roughly matches the distribution of policy implementation dates across California. Also, none of the stores are in jurisdictions that implemented policies before 2012, which means I have a full year of 2011 data in the pre-period for all stores.

The necessary identifying assumption for an event study design is that treated and control stores have parallel trends in the outcome variable pre-policy. Having stores that are also well matched on observables increases confidence that this assumption is satisfied. The top panel of Table 1 presents average store characteristics for treated and control stores. None of the variables are statistically different between treated and control groups. On average, stores in my sample first opened in 1985 and were remodeled in 2005. The majority of stores have bakery, deli, and floral departments, over half of the stores have pharmacies and coffee bars, and 10% or less have gas stations, juice bars, and sandwich counters. Roughly 45% of the stores (both treated and control) have self-checkout registers.

The bottom panel of Table 1 presents the summary statistics for average store demographics. Once again, none of the variables are statistically different across treatment groups.

[21] Of the 7 dropped stores, 2 store closed before the end of the sample period, 4 stores were remodeled to add self-checkout registers, and 1 store was in a jurisdiction where the DCB policy differed from the others in the sample in that it did not require a fee for paper bags. Given 5 of the dropped stores were in the treated group and 2 were in the control group, I lose 13% of the treated stores and 9% of the control stores.

Table 1 also presents the average demographics for California and for the United States. Comparing columns (1) and (2) with column (4), the stores in my sample are in areas with higher median incomes and a greater share of white residents than the California averages. These differences reflect that fact that DCB policy adoption occurred first in coastal California regions, which are more affluent on average than the Central Valley.

Finally, due to the constraints in obtaining data from the retailer, the sample includes only the hours between 1:00pm and 4:00pm for every Saturday and Sunday during the sample years. I chose these six weekend afternoon hours because they are among the peak shopping hours for supermarkets, which span 9:00am-5:00pm on weekends and 3:00pm-6:00pm on weekdays.[22] Having peak hours assures that transactions in the scanner data occur back-to-back, with little or no downtime in between. Moreover, peak hours, by definition, capture a large share of transactions. An industry white paper finds that half of grocery shopping transactions in the U.S. occur during peak hours, where a peak hour is defined as a time wherein more than 3 million people shop during that hour of the week (Goodman, 2008). In total, my sample of California scanner data has 1,047 hours of data across 53 stores and 3.5 years, for approximately 133 million items scanned and 9.8 million transactions.[23]

2.3 Outcome Variables in Scanner Data

Each observation in the scanner dataset corresponds to a purchased item, which I group into checkout transactions using a transaction identifier. For each item purchased within each transaction, the scanner data includes information on the item's name, Universal Product Code, and the purchase price. For each checkout transaction, the data include the time and

[22] Peak shopping hours for the stores in the sample were confirmed by the retailer providing the scanner data and verified using Google store hour data.

[23] I drop December 25 from the sample as not all stores are open on Christmas. I also drop Super Bowl Sundays as shopping patterns differ greatly on these days. Finally, I drop 142 transactions with more than 200 items scanned, as these were outliers.

date the transaction completed, the store identifier, the checkout lane number, a masked cashier identifier, and a masked customer card identifier. Using the identifiers, I am able to track stores, as well as cashiers and a subset of customers that frequently use rewards cards, over time.[24]

My main outcome variable for pre- and post-policy comparisons is *Transaction Duration*—the duration of each checkout transaction measured in minutes, from the start of a transaction until the start of the next transaction in line. I am able to construct this variable using the transaction time-stamp, which includes the day, hour, and minute each transaction was completed. Since I only have one time-stamp per transaction, I designed the sample to include only peak hours partly in order to make the assumption that transactions occur back-to-back, with little or no downtime in between. In Section 5, I verify this assumption using observational data collected in-store during peak hours, where transaction duration is timed with a stopwatch by enumerators stationed near checkout.[25] My second outcome variable is *Transactions per Shift*—the number of transactions completed in a store per 1:00-4:00pm weekend shift.

The top half of Table 2 presents transaction-level summary statistics for 2011, which predate all DCB policies in my sample (i.e., I have a full year of pre-policy data for all stores). Transactions are separated by the register type in which they occurred—1) full-service, 2) express, and 3) self-checkout.[26] Overall, Table 2 indicates that treated and control

[24] For the main analysis (Section 4), I use panel data averaged to the store level. In the sensitivity analyses (Section 5.3), I use panel data averaged to the cashier and customer level.

[25] To ensure that transactions occur back-to-back, I drop all transactions that are more than three standard deviations longer than the average transaction of its size (in terms of number of items scanned) and all transactions that are longer than 20 minutes. In cleaning the data this way, I lose 1.71 percent of transactions.

[26] Express registers have prominent signs overhead requesting shoppers to limit transactions to 15 items or fewer. Full-service registers have no recommended item limit. Self-checkout registers are registers where shoppers scan and bag their own items. I do not include transactions at specialty registers (e.g., registers in customer service, deli, and bakery departments) because there are few of these transactions and they rarely occur back-to-back.

stores have balanced transaction-level characteristics in the pre-period. At treated stores, the average transaction at a full-service register takes 1.99 minutes to complete, comprises of 19.22 items, and costs \$56.38.[27] The average transaction at an express register takes 1.48 minutes to complete, comprises of 8.57 items, and costs \$25.89. Finally, the average transaction at self-checkout registers takes longer to complete, contains fewer items, and costs less than at either full-service or express registers.

The bottom half of Table 2 reports average store-shift characteristics in 2011 for treated and control stores. In the pre-policy period, treated stores process approximately 567 transactions, \$21,811 in sales, and 7,380 items per 1:00-4:00pm weekend shift. Table 2 also presents the number of registers open on average and the total register capacity. To calculate the average number of registers open, I count the number of registers reporting at least one transaction per hour interval during the 1:00-4:00pm shift. Comparing the average number of registers open to the stores' register capacity, I find that stores are operating close to their full register capacity, at 2 fewer registers open than capacity on average. This suggests that during the peak weekend hours of 1:00-4:00pm, stores may be constrained by their fixed checkout capital. This will limit in the short run how stores can react to increases in checkout duration and congestion due to a policy shock.

2.4 Addressing Limitations of Scanner Data

While few datasets are as rich as the scanner data with respect to transaction-level knowledge, the sample design—by being constrained to weekend peak hours—presents potential concerns about whether the results are externally valid in all hours of the week. To address these concerns, I employ two additional datasets: 1) retail data measuring total sales at the

[27] The amount paid is created by summing up the individual amounts paid per item in a transaction. This variable does not include sales tax. Furthermore, several line items, including the line item for purchasing a paper bag and for making a donations to charity, do not include an amount paid.

store-month level, and 2) scanner data from weekday hours. The first dataset—collected by AC Nielsen and described in Section 6.1—consists of weekly price and quantity information at the product level for participating retail chains across the U.S.[28] Aggregating across all products sold in a store per week, these data can be used to track total store sales over time, and importantly, are not limited to certain hours in a week. The second dataset comes from the same chain as the main analysis, but from the District of Columbia (DC) metropolitan area. This scanner dataset—described in Section 7.3—was designed to study the effects of the 2010 DC bag tax. By spanning a shorter window of months than the California sample,[29] these data were able to include more hours per week, which allows for the comparison of policy effects across weekday and weekend hours. Together, these additional datasets provide strong evidence that the policy effects are not unique to peak weekend hours.

3 Empirical Model: Event Study Design

I estimate the causal effect of DCB policies on checkout duration during peak hours using an event study design. This approach can be thought of as unpacking a difference-in-differences (DID) design. Since each treated store can have a unique pre-/post-period, the event study model reorders the panel data to align the treatment events so that the differences in outcomes between treated and control stores can be plotted over event-time.

I average the transaction-level scanner data to one observation per store per week and employ the following event study regression model:

$$Y_{sjw} = \sum_{l=-24}^{24} \beta_l D_{l,jw} + \beta_x X_{sjw} + \theta_{sj} + \delta_w + \epsilon_{sjw} \tag{1}$$

[28]These data are made available through the Kilts Center at The University of Chicago Booth School of Business.

[29]The DC scanner dataset covers 4 months in the pre-tax period (Dec. 2008, Jan. 2009, Feb. 2009, Dec. 2009) and 5 months in the post-tax period (Jan. 2010, Feb. 2010 Dec. 2010, Jan. 2011 and Feb. 2011).

where Y_{sjw} is the outcome variable for store s in jurisdiction j and week-of-sample w, X_{sjw} is a set of control variables, θ_{sj} is a vector of store fixed effects, and δ_w is a vector of week-of-sample fixed effects. $D_{l,jw}$ is a dummy variable equaling one if jurisdiction j in week w implemented a DCB policy l weeks ago, with $l = 0$ denoting the week of implementation. The endpoints are binned, with $D_{24,jw} = 1$ for all weeks in which it is 24 weeks or more since DCB policy implementation and, similarly, $D_{-24,jw} = 1$ for all weeks in which it is 24 weeks or more until implementation.[30] The week prior to implementation ($l = -1$) is the omitted category. Store fixed effects control for time-invariant store level characteristics (i.e., store size, number of registers, types of departments offered). Week-of-sample fixed effects control for variation over time that effect all stores (i.e., holidays and seasons).

The β_l vector is the parameter of interest, as it traces out the adjustment path from before the DCB policies to after. I hypothesize that customer, cashier, and store learning will result in more complex dynamics than a simple discrete shift in the outcome variable (as would be implied by a model that replaced the $D_{l,jw}$ variables with a single indicator variable for the post-policy period). Customers must learn how to respond to the policy and change their habits (i.e., bring more bags from home, buy paper or reusable bags at checkout). Cashiers must alter their checkout procedures. Store managers may reoptimize the number of lanes open and the placement of cashiers as to keep lines to a minimum. All of these behaviors may change over time as customers, cashiers, and stores learn, adapt, and circumvent the new policies.

Therefore, I expect that the effects of the policy will be greater in the initial weeks, and

[30] I choose ±24 weeks as endpoints because I hypothesize that 24 weeks (or roughly half a year) is enough time to witness learning. I also bin at +24 weeks because stores that implement policies later in the sample period mechanically have fewer post-policy weeks than stores with early implementation dates. While all 33 treated stores have at least fifteen weeks in the post-policy period, only 28 stores have thirty weeks, only 23 stores have sixty weeks, only 13 stores have eighty weeks, and so on. Thus, binning the endpoints at 24 weeks provides ample time for measuring learning without losing too many of the treated stores. I will also examine whether the results are robust to binning at -48 and +96 weeks.

will diminish over time (i.e., β_0 will be greater in magnitude than β_{24}). To test this formally, I will use two Wald tests. In the first test, the null hypothesis is that all coefficients in the post-policy are equal (i.e., $\beta_0 = \beta_1 = \beta_2 = ... = \beta_{24}$) and in the second test, the null hypothesis is that the coefficient for the first week of the policy is equal to the coefficient for all weeks 24 or more after the policy (i.e., $\beta_0 = \beta_{24}$). Rejecting these hypotheses would provide evidence of learning.

The identifying assumption of the model is that, absent the DCB policies, outcomes at the treated stores would have remained similar to the control stores. Underlying trends in the outcome variable correlated with DCB policy enactment are the most likely violation of this assumption. Part of the appeal of event study designs is that the pre-policy portion of the β_l vector provides a check against this possible violation. If DCB policies are unassociated with underlying trends, there should be no trend in the β_l vector in the pre-policy period (i.e., $\beta_{-24} = \beta_{-23} = \beta_{-22} = ... = \beta_{-1}$).

The primary outcome variables I use for Y_{sjw} will be (1) logged average transaction duration, measured in minutes, and (2) average number of transactions completed per 1:00-4:00pm shift. I examine additional outcome variables as well, such as average share of transactions purchasing paper and reusable bags, and the number of registers open.

4 Result 1: Effect of Bag Bans on Transaction Duration

4.1 Average Effects

The figures in this section present the results from the estimation of event study Equation 1, where the $\hat{\beta}_l$ point estimates and 95% confidence intervals are displayed graphically.[31] Unless specified otherwise, I cluster the standard errors two ways—by jurisdiction (45) and by week-

[31] I estimate all fixed-effect equations in STATA using the command `reghdfe` (Correia, 2014).

of-sample (177)—to allow for spatial and temporal correlation in the data.[32]

In Figure 3, the transaction-level scanner data are averaged to the store-week level, for a total of 9,381 observations. The outcome variable, Y_{sjw}, is logged average transaction duration, which means the $\hat{\beta}_l$ point estimates measure the percent difference in transaction duration between control and treated stores l weeks from policy implementation. Panel (a) displays the results for the simplest specification, which includes the event study indicators, store fixed effects, and week-of-sample fixed effects. Variations in grocery shopping demand by store and week-of-sample (such as from local festivals and sporting events) may influence checkout duration, and these variations are not absorbed by the store and week-of-sample fixed effects. To account for grocery shopping demand which varies by store and week, the specification in panel (b) additionally includes control variables, X_{sjw}, for the average number of items purchased per transaction, the average dollar amount spent per transaction, and the share of transactions purchasing (a) produce, (b) fresh meat and seafood, (c) dairy and refrigerated, (d) frozen, (e) bakery and deli, (f) shelf-stable food, (g) alcohol and tobacco, (h) infant/toddler, (i) floral department, and (j) pet items. In Section 5, I verify that these control variables are not bad controls—i.e., the average number and types of items purchased does not change with the implementation of DCB policies.

In panel (a), I find strong evidence that the DCB policies lead to increased average transaction duration. The slowdown starts in the first week of the policy with $\hat{\beta}_0 = 0.037$, which means the average transaction duration at treated stores is 3.7% longer during the first week of a DCB policy. The slowdown fluctuates slightly over time, peaking with $\hat{\beta}_4 = 0.048$ and ending with $\hat{\beta}_{24} = 0.025$. The $\hat{\beta}_{24}$ coefficient indicates that for all weeks in which it has been 24 or more weeks since policy implementation, transactions at treated stores

[32] Estimating a model that allows for spatial correlation up to 12 km and temporal correlation up to 8 weeks using spatial errors—as described by Conley (2008) and implemented using code from Hsiang (2010) and Fetzer (2014)—does not change the significance of the results.

remain 2.5% longer than at control stores.[33] The majority of the post-policy $\hat{\beta}_l$ coefficients are significantly greater than zero at the 10% significance level. Importantly, only one of the pre-policy $\hat{\beta}_l$ coefficients is significantly different from zero, which provides evidence in favor of the identifying assumption that transaction durations at treated stores were not trending differently than at control stores before the DCB policies went into effect. Panel (b) shows that the inclusion of control variables does not greatly alter the $\hat{\beta}_l$ coefficients. Unless otherwise specified, I will use the full model specification with control variables in the remainder of the paper.

Using Wald tests to compare the event study coefficients in Figure 3b, I can reject that all $\hat{\beta}_l$ coefficients in the post-policy period are equal, however, I cannot reject that $\hat{\beta}_0 = \hat{\beta}_{24}$. These results suggest that DCB policies lead to a persistent increase in transaction duration over the sample period relative to control stores. In other words, with data aggregated to the store-week, I do not find evidence of transaction durations returning to pre-policy levels over time.

A potential concern is that 24 weeks (or roughly half a year) is not enough time to witness learning. In Figure 4, I explore whether the effects of DCB policies on transaction duration lessen over time if the event study model is binned at –48 and +96 weeks (i.e., roughly 1 year before and 2 years after) instead of ±24 weeks. I find the 3% slowdown in transactions duration persists even when the event study is binned at –48 and +96 weeks. However, the $\hat{\beta}_l$ estimates grow noisier after D_{73}, when the number of treated stores in the sample drops to 13.

In addition to the event study model in Equation 1, I estimate the following DID model:

$$Y_{sjw} = \beta_D D_{jw} + \beta_x X_{sjw} + \theta_{sj} + \delta_w + \epsilon_{sjw} \tag{2}$$

[33]The numerical regression output for Figure 3a can be found in Online Appendix Table A.2.

where D_{jw} is now a single dummy variable equal to 1 when a DCB policy is in effect in jurisdiction j and week-of-sample w, instead of the set of event study dummy variables. The results are presented in column (1) of Table 3. I estimate $\hat{\beta}_D = 0.033$ (p-value = 0.000), which is consistent with the event study results in Figure 3.

4.2 Heterogeneity by Bag Choice

Given that I find a statistically significant and persistent 3% slowdown in transaction duration due to DCB policies on average, I next investigate mechanisms behind the slowdown, and in particular, whether the effects of the policies are heterogeneous by the types of bag chosen. First, I estimate Equation 1 with Y_{sjw} being (i) the share of transactions paying for at least one paper bag in the post-policy period and (ii) the share of transactions purchasing at least one reusable bag.[34] Figure 5 presents the results.

As one would expect, I find a sharp and permanent increase in the share of customers purchasing paper bags which is contemporaneous with DCB policy implementation (panel a). Since paper bags were available but not sold before the policies, this figure reassures me that I have the correct timing of the policy implementation. Approximately 27% of transactions chose to purchase paper bags in the first week of the policy. This share drops 5 percentage points in the second and third weeks of the policy and remains at roughly 20% for the remainder of the sample period. For reusable bags in panel (b), I find a temporary increase in purchases of reusable bags when the DCB policies are implemented. Reusable bags are sold at the supermarket chain both before and after the policies.[35] As with paper bags, the share of transactions choosing to buy reusable bags increases sharply at the onset of DCB

[34] In the scanner data I see whether or not a transaction pays the paper bag fee, but not how many paper bags are purchased.

[35] The prices of reusable bags do not differ between treated and control stores and they also do not change when the DCB policies go into effect. I find this both in the scanner data and during in-store visits.

policies, with 5% of transactions choosing to purchase reusable bags in the first and second weeks of the policy. However, this increase quickly retreats, and by eight weeks after the policy implementation the share of transactions purchasing reusable bags is indistinguishable from zero. This pattern is consistent with customers reusing the reusable bags they purchase in the first weeks of the policy.

Given the persistent increase in paper bag use, I next explore how customer bag choice influences the effects of DCB policies on transaction duration. I estimate Equation 1 by whether a transaction purchased a paper bag in the post-policy period (i.e. the transaction-level scanner data are averaged to the store-by-week level for those that purchase paper at treated stores in the post-policy period and for those that do not). Figure 6 presents the results. I find a stark difference in policy effects between transactions with and without paper bag purchases, with transactions paying for paper bags experiencing more than three times larger slowdowns than transactions forgoing paper bags. For instance, $\hat{\beta}_0$ is equal to 0.017 for transactions not paying for paper bags (panel a), while it is equal to 0.092 for transaction paying for paper bags (panel b). These results provide evidence that bag type is an important mechanism behind the persistent slowdown, with the slowdown loading on the transactions that pay for paper bags.

It should be noted that these results are not identifying a causal effect of choosing paper bags on transaction duration, as I do not randomly assign who gets paper and who does not. Customers who choose to pay for paper bags could be inherently slower than those that do not.[36] Furthermore, splitting the transactions by whether a paper bag was purchased in the post-period means that the treated customers in the pre-period are not necessarily the same as the treated customers in the post-period. To alleviate these concerns, I use customer-

[36] At the transaction level, paper bag use is positively correlated with transaction size and purchasing more expensive items.

level data to split customers at treated stores in two groups: those that never purchase a paper bag and those that ever purchase a paper bag (presented in Online Appendix B). I find that neither of the treated customer groups differ from the control customers in the pre-period. Yet after the policy, treated customers that choose paper bags have longer transaction durations than those who never purchase paper, relative to the control customers. Section 7 further tests the robustness of this mechanism. First, I use data collected in-store to verify whether and by how much paper bags are a slower technology than plastic bags. Second, I use scanner data from stores affected by a plastic bag tax instead of a plastic bag ban to test whether DCB policies lead to slower transactions when plastic bags—not paper bags—remain the default option.

4.3 The Time Cost of DCB Policies

What are the implications of a 3% slowdown? A 3% increase in transaction duration means that the average 2 minute transaction is approximately 3.6 seconds slower. While 3.6 seconds might seem like a trivial amount of time, when aggregated across all shopping trips made per year, this time cost becomes substantial. In the United States, an estimated 11.9 billion grocery shopping trips are made annually,[37] meaning 3.6 seconds per grocery shopping trip would aggregate to 11.9 million additional grocery shopping hours per year. Using half the average U.S. hourly wage as the value of time—since grocery shopping often occurs during non-work hours when the opportunity cost of time is low[38]—11.9 million hours is worth $138.3 million. This back-of-the-envelope calculation relies on the assumption that

[37] Hamrick et al. (2011) estimate how much time Americans spend on food and find the average U.S. adult grocery shops once every 7.2 days. Given roughly 235 million adults in the U.S. (2010 U.S. Census), this equates to 11.9 billion grocery shopping trips per year.

[38] I use half the hourly wage because it is a generally accepted figure for the value of non-work time (Small, 1992; Small and Verhoef, 2007). Half the average U.S. hourly wage was $11.62 in 2015 ("May 2015 State Occupational Employment and Wage Estimates." *Bureau of Labor Statistics, U.S. Dept. of Labor* Online, *accessed March 12, 2018*).

transactions during non-peak hours—which are roughly half of all transactions[39]—experience the same slowdown from the DCB policies as transactions during peak hours. In particular, if the share of paper bag use is more (less) during off-peak hours, the average slowdown during off-peak hours could be higher (lower).

Another reason 3.6 seconds per transaction is not trivial is because checkout transactions are connected though queuing systems. During peak hours, when stores are operating at their full register capacity and checkout transactions occur back-to-back, a customer not only has to wait the extra time for their own transaction, they must also wait the extra time for all the customers ahead of them in line. Even though the scanner data does not measure queue length directly, a simple queuing model can be used to approximate how much longer customers would wait in line due to DCB policies.[40] Calibrating this model with the scanner data demonstrates that a 3.6 second increase in every transaction during peak hours would compound so that the average customer checking out during these hours would wait an additional 104 seconds. Based on industry reports, this is a 43% increase in the average wait time of supermarket checkout.[41] This congestion externality would be less if stores open more lanes or if the arrival rate of customers decreases. In sections 5.2 and 6.1, I explore whether DCB policies cause (1) stores to open more registers and (2) customers to shop less frequently at affected stores.

To get a sense of the magnitude and relevance of this time cost, one can look at another recent change to cause longer checkout wait time—chip technology to reduce credit card

[39] An industry white paper finds that half of grocery shopping transactions in the U.S. occur during peak hours (Goodman, 2008), where a peak hour is defined as a time wherein more than 3 million people shop during that hour of the week.

[40] See Online Appendix A for a full discussion of the queuing model and assumptions.

[41] An industry survey found that the average wait time in grocery shopping lines in 25 major cities was 4 minutes. Source: "Justice—Wait for It—on the Checkout Line." *Wall Street Journal*. Aug. 19, 2009. Online, *accessed May 30, 2016.*

fraud.[42] An industry study found that using a chip card added 8 to 12 seconds per checkout transaction.[43] Similar to paper bags being slower to pack than plastic bags, these slowdowns come from chip readers having slower software than swipe readers. After months of customer consternation and retailers choosing to be liable for fraudulent charges over activating chip readers, Visa and MasterCard released software to bring down the processing time of their cards and the US's largest retailer, Wal-Mart, eliminated machine prompts to shave seconds off.[43] The experience of chip cards suggests customers and retailers are aware of and sensitive to the extra seconds spent in checkout.

5 Alternative Mechanisms Behind Slowdown

In the event study results presented above, I find that DCB policies lead to statistically significant and persistent increases in transaction duration. Additionally, I find that the effects of DCB policies are greater for transactions paying the paper bag fee. In this section I rule out three alternative mechanisms for why DCB policies lead to checkout slowdowns.

5.1 Mechanism 1: Do DCB policies alter how much customers purchase?

Are the slowdowns in transaction duration driven by changes in what customers buy when the DCB policies go into effect? In Figure 7, I examine whether the number of items purchased and the amount spent per transaction changes when DCB policies are implemented. I estimate the simplest specification of Equation 1, with only store and week-of-sample fixed effects. Each panel of Figure 7 has a different outcome variables, Y_{sjw}, at the store-week

[42] On Oct. 1, 2015, retailers that did not implement chip payment terminals would face liability for fraudulent charges in their stores for which banks and payment processors were previously liable. As of December 31, 2015, only 20% of retailers had activated terminals and an additional 30% had terminals installed but not activated. Conversely, almost 60% of credit cards issued by banks were embedded with a chip. ("Chip Cards Cause Headaches at Stores Across America." *Bloomberg.* Apr. 13, 2016. Online, *accessed Jul. 22, 2016.*)

[43] "Visa, Wal-Mart Move to Speed Checkout for Customers with Chip-enable Cards." *Wall Street Journal.* Apr. 19, 2016. Online, *accessed Jul. 22, 2016.*

level: (a) the average number of items scanned per transaction, not including checkout bags, (b) the average dollars spent per transaction, and (c) the average dollars spent per item.[44]

In panels (a) and (b), I do not find evidence that DCB policies lead to changes in the average number of items purchased or in the average amount spent per transaction, with the majority of $\hat{\beta}_l$ coefficients statistically indistinguishable from zero. In panel (c), I find some evidence of a temporary dip in the average amount spent per item. Specifically, in the second week of the policy, the average amount spent per item is 2.4 cents lower than at control stores ($\hat{\beta}_1 = -0.024$). While I do not observe the size of items purchased (i.e., 8oz vs 20oz), this is consistent with the hypothesis that DCB policies alter the size of the items purchased, with customers preferring smaller (and less expensive) items when they need to pay for, or remember, checkout bags. However, since this change is temporary and quite small in magnitude,[45] it is unlikely to be the mechanism behind the persistent slowdown in checkout transactions. The results in all three panels are the same if I use median values instead of mean values—further evidence that DCB policies do not shift the distribution of transaction size.

I also estimate Equations 1 and 2 with the outcome variable being the share of transactions in store s, jurisdiction j, and week-of-sample w purchasing items in the following categories: 1) produce, 2) meat and seafood, 3) dairy and refrigerated, 4) frozen, 5) bakery and deli, 6) shelf-stable food, 7) alcohol and tobacco, 8) baby, 9) floral, and 10) pet. Except for produce, I find no significant changes due to the DCB policies for these items. While I find a 0.4 percentage point decrease in the share of transactions purchasing produce (statis-

[44] The dollars spent variable is created by summing up the individual amounts spent per item in a transaction, and therefore, it does not include sales tax. Several point of sale line items, including the line item for purchasing a paper bag and for making a donation to charity, do not include an amount spent. Since the amount spent variable does not include paper bags purchased, I measure the average amount paid per item as the amount paid per transaction divided by the number of items scanned not including checkout bags.

[45] The average item in 2011 costs \$3.01, so a 2.4 cent drop in price is less than a 1% change.

tically significant at the 10% level), given that 52 percent of transactions in the pre-policy period purchase produce, this is a relatively small change.[46]

5.2 Mechanism 2: Do DCB policies alter which registers customers use?

Along with choosing how many and what type of groceries to buy, customers also choose at which register to queue. In Figure 8, I estimate Equation 1 with the outcome variable being the share of transactions in cashier-operated registers versus self-checkout registers. In panels (a), I find that after the DCB policies are implemented, the share of transactions completed at cashier-operated registers declines slightly over time relative to self-checkout registers. In particular, $\hat{\beta}_{24} = -0.013$ in panel (a), indicating a roughly 1 percentage point decrease in the share of transactions at cashier-operated lanes 24 weeks after policy implementation.

This result suggests that some customers adapt to the policy by switching from full-service to self-checkout lanes. Adopting a new technology, such as self-checkout, is often spurred by dramatic events that change the effort and time of the alternatives. While transaction duration at self-checkout registers are on average 2 minutes longer than full-service transactions (as seen in Table 2), the self-checkout queues may be relatively shorter after the DCB policies, inducing customers to switch. Learning-by-doing might also be at play. The DCB policies may lead customers to try self-checkout registers for the first time, and having used the self-checkout once, they are more likely to do so in the future. Finally, bringing ones own bags to the store may change consumers' preferences over having other people bag their groceries. Yet, while the increased use of the slower self-checkout technology may explain some of the persistent effects of DCB policies on transaction duration, it cannot explain the initial slowdown.

Second, to see whether stores open more registers in response to the DCB policies, I also

[46] The DID estimates can be found in Online Appendix Table A.3.

estimate Equation 1 with the outcome variable being the average number of cashier-operated and self-checkout registers open. In panel (b), I find that stores temporarily increased the number of cashier-operated registers open. In the second week of the policy, stores had 0.3 more registers operating on average. By nine weeks after the policy, stores were operating the same number of registers as before the policies. In panel (c), I find no statistically significant changes in the number of self-checkout registers open, which is not surprising given self-checkout registers are turned-on as long as they are not out-of-order. Therefore, it appears that stores brought on more staff during the initial weeks of the policy to mitigate the effects of DCB policies and assist with the transition. In Section 7, I examine whether stores alter other operation behaviors—such as the number of baggers present—using data collected in-store.

5.3 Mechanism 3: Do changes in the composition of cashiers and customers drive the results?

In the store-week events studies presented in Section 4, I use data averaged to the store-week level in order to eliminate concerns over correlation between transactions within a store and week leading to inconsistent standard errors.[47] However, given the high turnover of cashiers and the heterogeneity of customers, using store-week data may hide changes in the composition of cashiers and customers, and these compositional changes could be an alternative mechanism behind the slowdown. In this section I explore the sensitivity of my results to estimating the model at the cashier level, with cashier fixed effects. In Online Appendix B, I similarly explore the sensitivity of the results at the customer level,

[47] Bertrand et al. (2004) discuss issues with estimating DID regressions, and find that when more than two periods of data are used, there is a potential for a large number of dependent observations within each cross-sectional unit. One of the solutions they test and recommend is to collapse the data until the dependence issue disappears.

with customer fixed effects.[48]

Supermarket cashier is a position with high-turnover, and the cashiers present at the beginning of the sample are not that same as those at the end. Thus, I average the transaction-level data to the cashier-week level and examine whether including cashier fixed effects alters the results. The event study model at the cashier-week level is as follows:

$$Y_{csjw} = \sum_{l=-24}^{24} \beta_l D_{l,jw} + \beta_x X_{csjw} + \alpha_{csj} + \delta_w + \epsilon_{csjw} \tag{3}$$

where Equation 3 uses average data for cashier c at store s, jurisdiction j, and week-of-sample w. Importantly, the inclusion of cashier fixed effects, α_{csj}, means the β_l coefficients in Equation 3 measure the policy effects *within cashiers* over time.

Figure 9 presents the cashier-week event study results from estimating Equation 3, where I display the $\hat{\beta}_l$ point estimates and standard errors graphically. The outcome variable, Y_{csjw}, is the logged average transaction duration for cashier c at store s, in jurisdiction j and week-of-sample w. In addition to cashier and week-of-sample fixed effects, I control for the average number of items scanned and amount spent per transaction for cashier c in week-of-sample w, as well as the types of items purchased. I also control for the experience of cashiers, using indicator variables for the number of weeks cashier c had worked the 1:00-4:00pm shift in store s and week-of-sample w.[49]

Figure 9 presents a slightly different pattern than the store-week analysis in Figure 3.

[48] The important take-away from the customer level analysis presented in Online Appendix B, is that including customer fixed effects does not alter the main results in Figure 3. I find that DCB policies lead to sharp and persistent increases in checkout duration *within customer*, which mean the checkout slowdown in the store level analysis is not driven by changes in the composition of customers.

[49] I drop cashiers who are in the sample fewer than 18 weeks (or roughly 4 months), in order to have cashiers that are in the sample long enough to experience learning. This gives me a total of 1,914 cashiers across the 53 store. On average, 36 cashiers work the 1:00-4:00pm weekend shift per store over the 3.5 years of the sample, with a minimum of 19 cashiers and a max of 52 cashiers per store. The median number of weeks worked by cashiers during the sample is 49 out of 177.

First, at the cashier-week level I find that the slowdown in transaction duration began a week or two before the policy.[50] When interviewed, store managers explained that they took measures to prepare their stores for the policy change in the weeks before implementation. In particular, cashiers were asked to start reminding customers of the upcoming policy change. Second, I find that the initial slowdown during the first week of the policy is greater at the cashier level than at the store level. In Figure 9, $\hat{\beta}_0 = 0.054$, while in Figure 3b, $\hat{\beta}_0 = 0.040$. Third, there is stronger evidence of learning at the cashier level than at the store level, with the post-policy $\hat{\beta}_l$ coefficients diminishing in size over time ($\hat{\beta}_{24} = 0.026$). Using Wald tests to compare the coefficients, I can reject that all $\hat{\beta}_l$ coefficients in the post-policy period are the same as one another at a 1% significance level, and I can reject that $\hat{\beta}_0 = \hat{\beta}_{24}$ at the 10% significance level.

Figure 9 suggests that cashiers do learn and get faster after DCB policies; however, cashier learning of 2-3% does not completely offset the 5% initial within cashier slowdown from the policies. Thus the reduction in productivity from DCB policies persists even after cashiers learn and adapt to the change.

6 Result 2: Effect of Bag Bans on Supermarket Sales & Shopping Frequency

6.1 Do DCB Policies Disincentivize Grocery Shopping?

The results above indicate that the DCB policies cause a significant and persistent 3% increase in transaction duration for the average customer. How does the time cost of DCB policies affect the likelihood that customers continue to shop at regulated supermarkets? More generally, how do time costs affect economic behavior? Since Becker (1965) first stressed the value of time in choice analysis, studies have shown that consumers dislike and

[50] In Figure 9, the omitted event study dummy is $D_{-2,jw}$ instead of $D_{-1,jw}$, so that the slowdown in the week before the policy is clearly visible.

actively avoid long wait times (Katz et al., 1991; Tom and Lucey, 1995; Van Riel et al., 2012). For instance, Craig et al. (2016) study the relationship between the length of time a blood donor waits to give blood and the likelihood that they return to give blood again. The authors find that a one standard deviation increase in the average wait results in a 10% decrease in blood donations per year. In a similar vein, I next explore how increases in checkout duration from DCB policies alter the likelihood of return visits and supermarket sales.

First, I quantify the change in the number of transactions completed per store and shift by estimating DID Equation 2 with the outcome variable being the average number of transactions completed per three-hour weekend shift. Table 3 presents the $\hat{\beta}_D$ coefficients estimated in both levels and logs. In column (2), I find that stores process 18.226 fewer transactions per three-hour shift when DCB policies go into effect. In Figure 10, I instead estimate event study Equation 1. Unlike the event study results using transaction duration as the outcome variable (Figure 3), I do not find an immediate change in the number of transactions processed, with β_0 indistinguishable from zero. However, overtime there is a statistically significant drop in the number of transactions processed, with $\beta_{24} = -21.214$.

Losing 21 transactions per peak hour shift might not be concerning if customers are simply shifting into less busy hours of the day. For the 53-store sample, I only have weekends transactions between 1:00-4:00pm, and consequently, I cannot measure directly from this dataset whether total weekly store sales decreased due to the DCB policies. However, other scanner datasets exist that track total sales at individual stores over time. Using a sample of the Nielsen Retail Data, containing 539 stores across 18 California jurisdictions from 2009 until 2015,[51] I estimate the following event study model:

$$(4) \qquad lnSales_{sjm} = \sum_{l=-12}^{12} \beta_l D_{l,jm} + \theta_{sj} + \delta_m + \epsilon_{sjm}$$

where $lnSales_{sjm}$ is logged sales occurring in store s, jurisdiction j, and month-of-sample m, $D_{l,jm}$ are the event study dummies, and θ_{sj} and δ_m are store and month-of-sample fixed effects. Figure 11 presents the results, with the outcome variable being logged sales either (a) measured in dollars or (b) measured in items sold. In panel (a), DCB policies lead to a drop in sales that increases over the first couple of months. In particular, there are 0.9% fewer sales in the first month of the policy ($\beta_0 = -0.009$), 1.9% fewer sales in fourth month of the policy ($\beta_3 = -0.019$), and 1.6% fewer sales 12 months or more after the policy ($\beta_{12} = -0.016$). A similar pattern holds in panel (b) for items sold per month. Together with the results above, this suggests that the transactions lost due to DCB policies during peak hours are not fully made up during the less busy hours of the week. In fact, stores in treated juridictions experience a 1.6% decrease in sales persistently a year and more after the policy. For the average store in the Nielsen sample, this is a $213,407 reduction in sales per year. If the Nielsen sample is representative of the 10,935 supermarkets and grocery stores in California,[52] a statewide DCB policy would cause a $2.3 billion loss in grocery sales annually.

Assuming people are still eating the same amount of food as before, where are these lost grocery sales going? Are they moving to online food purchases? Are they switching to fast food purchases and other types of restaurants? Not only is grocery shopping less convenient under DCB policies, restaurants and food-away-from-home establishments are

[51] For information on the Nielsen Retail Data and sample selection criteria, please see Online Appendix C.

[52] The number of supermarkets and other grocery stores in California comes from the California State Board of Equilization, *Taxable Sales in California 2014* Report.

often exempted from these policies and thus can still offer disposable plastic bags to their customers. Since dining out and grocery shopping are substitute goods, and DCB policies effectively raise the cost of grocery shopping relative to dining out (whether due to time costs or the actual costs of bags), some customers may have chosen to purchase food elsewhere. Online Appendix D presents an analysis of quarterly, county-level taxable sales by retail category in California.[53] I document a 3.7% increase in the taxable sales from Food Services and Drinking Places (e.g., full-service restaurants, fast food restaurants, bars) that is contemporaneous with the implementation of DCB policies. None of the other retail categories examined experience statistically significant changes in taxable sales. While these data are limited in their aggregation and scope (i.e., most foods sold for at-home consumption are not taxed in California and thus taxable sales data cannot be used for supermarket food sales), they provide suggestive evidence that DCB policies shifted some customers towards eating out more often.

This unintended consequence is concerning because away-from-home meals tend to be higher in energy density, fat, and sodium and lower in fruits, vegetables, whole grains, and nutrients than at-home meals (Kant et al., 2015). In order to fully understand the health consequences of DCB policies, future work needs to further explore where the customers that are grocery shopping less often shift their food purchases.

6.2 Hazard Model Linking Transaction Duration to Lost Supermarket Sales

The above analysis shows that DCB policies lead to significant increases in transaction duration and significant decreases in supermarket sales. However, how much of the decrease in sales is due to the increased checkout time cost? Sales might also decrease if customers object to the DCB policies and shop less as a form of reactance (Just and Hanks, 2015), or

[53] Source: California State Board of Equilization, *Taxable Sales in California* Reports.

sales might decrease if the policies increase the stigma customers face for forgetting a reusable bag. Furthermore, in addition to increasing the amount of time to checkout, DCB policies may also increase the amount of time needed before entering the store (e.g., remembering to bring reusable bags) and after exiting the store (e.g., cleaning reusable bags that contained raw meat and unwashed produce), increasing the time cost of grocery shopping in other ways.

In order to disentangle the effect of longer checkout durations on the likelihood of return visits, the ideal dataset would include the number of days between customers' store visits linked to the processing time of each of their checkout transactions. These data partially exist in the 53-store sample for the 1,117,009 customers that use their customer cards more than once during peak weekend hours. For these customers, I know how long their transaction took to complete and how many weeks until their next peak hour visit. Using the sample of customer card users, averaged to the customer-week level, I estimate the following hazard model:

$$h(t|\mathbf{x}_{isjw}) = exp(\gamma t)exp(\beta_0 + \beta_1 D_{jw} + \beta_2 \widetilde{TxnDur}_{isjw} + \theta_{sj} + \delta_w) \tag{5}$$

where the baseline hazard model is parameterized using the Gompertz distribution with shape parameter γ.[54] The dependent variable is the probability of shopping during peak hours in week t, where t counts the number of weeks since the last visit. D_{jw} is again a dummy variable equal to 1 when a DCB policy is in place in jurisdiction j and week-of-sample w. I hypothesize that, ceteris paribus, DCB policies will have a negative effect on customers' likelihood to return in subsequent weeks. $\widetilde{TxnDur}_{isw}$ is the residual duration of

[54] I present a parameterized hazard model over a Cox proportional hazard model, which does not parameterize the baseline hazard, in order to estimate the expected number of weeks a customer delays their return. However, using a Cox proportional model instead does not alter the results, suggesting the results are not sensitive to parametrization.

customer i's transaction after regressing transaction duration ($TxnDur_{isw}$) on the number of items scanned, amount paid, and indicators for the types of items purchased and for whether the transaction occurred at a self-checkout register. This transformation to demean $TxnDur_{isw}$ is performed so that β_1 is interpreted as the relationship between the likelihood to return in subsequent weeks and deviations from the average duration of transactions with similar characteristics. I hypothesize that, ceteris paribus, longer than average transaction durations will have a negative effect on customers' future shopping behavior. That is, the longer than average a customer waits during checkout, the less likely they will return at time t conditional on having not yet returned. Finally, θ_{sj} and δ_w are fixed effects for store s in jurisdiction j and week-of-sample w.

Table 4 reports estimates of Equation 5. In all columns, the negative coefficients on *Ban Effective* and *Txn. Duration* support the hypotheses that shoppers who experience either a DCB policy or longer than average processing times are less likely to return during peak shopping hours in future weeks. Furthermore, not including *Txn. Duration* in column (2) increases the magnitude of the coefficient on *Ban Effective* from -0.009 to -0.010. This suggests one tenth of the reduction in the likelihood to return due to DCB policies is driven by the longer than average transaction durations that DCB policies cause. This is verified by evaluating the hazard model at the average increase in transaction duration due to the policies, which is 3.6 seconds or 0.06 minutes. Using the model in column (1), a shopper who experiences a DCB policy is 0.9% less likely to return during any given week and a shopper that experiences a transaction that is 0.06 minutes longer than average is 0.066% ($= 0.011 * 0.06$) less likely to return. Together these effects are similar to what is found in column (2) where transaction length is not included in the model (i.e., $0.966\% \approx 1.0\%$). It should be noted that this estimate is smaller than the estimate found using the Nielsen data, where DCB policies led to a 1.6% drop in sales. In column (3), not including *Ban*

Effective does not alter the estimates for *Txn. Duration* from column (1), and in column (4), including an interaction term between *Ban Effective* and *Txn. Duration* also does not alter the estimates from column (1). These results suggest DCB policies do not alter customers' dislike for longer transaction times.

One way to quantify the change in shopping behavior due to longer checkout durations is to consider the change in behavior of the median customer—the customer who returns after 50% of the other customers have already returned. Using the parameterization of the model in column (1) of Table 4, a 0.06 minute slowdown in checkout duration leads to 0.004 fewer store visits per year during peak hours for the median customer card user. If all 28 million adults in California have a similar sensitivity to checkout duration, this would aggregate to 112,000 fewer store visits in California per year just because of the persistent 3% increase in checkout duration. Therefore, while a permanent 3% increase in checkout duration may appear inconsequential, consumers respond to this time cost by shopping less often, and due to the pervasiveness of grocery shopping, this time cost and its resulting behavioral changes aggregate quickly.

In summary, a 3% increase in transaction duration can explain roughly 10% of the reduction in sales from DCB policies for the sample of customer card users shopping during peak hours. An obvious drawback of this analysis is that it can say little about customers that did not use customer cards or that shop at less busy hours of the week. However, these results show that some consumers are aware of and sensitive to time costs at checkout when deciding their shopping frequency.

7 Robustness and External Validity: Evidence from Supplementary Data

In the following subsections I explore supplementary datasets to test the robustness of the results above as well as their external validity. In subsection 7.1, I compare the effects of

DCB policies on transaction duration using scanner data versus using observational data collected in-store. In subsection 7.2, I estimate the effects of DCB policies at an alternative supermarket chain, to investigate whether the checkout slowdowns are a general phenomenon or unique to the retail chain in the main analysis. In subsection 7.3, I analyze whether slowdowns occur under a different type of policy (i.e., a bag tax), in a different state and time period.

7.1 Robustness Analysis 1: Scanner Data vs. In-store Data

While the supermarket scanner dataset is rich along several dimensions, it is missing three key variables: i) the presence of baggers at checkout, ii) the types and number of bags customers use before and after the policy change, both purchased and brought from home, and iii) the amount of downtime, if any, between transactions. To address these data limitations, I designed a follow-up field experiment—taking advantage a DCB policy implemented in Contra Costa County California on January 1, 2014. A team of enumerators made bi-weekly visits to three supermarkets—of the same retail supermarket chain as the scanner data—to collect data through direct observation of checkout transactions. The enumerators, stationed near full-service registers, collected information on the number and types of bags used, the presence of a bagger, the duration of each transaction, and basic customer demographic information such as gender and race of the person paying.[55] These visits were made over five months—one month before (December) and four months after (January-April) the policy change in Contra Costa County. Each visit lasted 1-2 hours and was made on either a Saturday or Sunday between 11:00am and 7:00pm. I also obtain the scanner data for the same dates and hours as the in-store visits. In this subsection, I use the in-store data to examine the effects of DCB policies controlling for variables that cannot be measured with

[55]Observations were made only at full-service registers, and not express or self-checkout registers.

the scanner data.

The first store, which I refer to as the *treated* store, is in Richmond, a city that implemented a DCB policy during my sample period. The second store, which I refer to as the *prior-policy* store, is in Berkeley, a city that adopted a DCB policy in January 1, 2013, exactly one year before the Richmond policy. The third store, which I refer to as the *no-policy* store, is in Concord, a city that had yet to adopt a DCB policy during the sample period. The two control cities were chosen to match Richmond with respect to average demographic characteristics.[56]

How do the in-store and scanner datasets compare along the variable of interest—transaction duration? In particular, I am concerned that my measure of transaction duration in the scanner dataset may overestimate the actual transaction duration because of the potential downtime in between transactions that is missing in the scanner dataset. In Table 5, I compare the average transaction duration, measured in minutes, for the in-store and scanner datasets. For the full sample of transaction, the average transaction duration in the in-store dataset is 0.119 minutes shorter than in the scanner dataset, which translates to roughly 7.14 seconds. Thus I do find that the scanner data misses a portion of downtime in between transactions. However, the worry is not that this difference occurs but that it happens differentially at stores with and without DCB policies. Thus I compare the average transaction duration between in-store and scanner for stores with and without DCB policies. Reassuringly, I find similar differences between in-store and scanner datasets when splitting the sample by policy treatment.

I next examine the effects of DCB policies on transaction duration at full-service registers

[56]The in-store data collection was designed to answer multiple questions about the effects of DCB policies. In Taylor and Villas-Boas (2016b), the in-store data were used to measure how checkout bag choices change when DCB policies go into effect. Please see Online Appendix E for a more detailed description of the variables in the in-store data.

using the in-store observational data. I estimate the following event study model:

$$(6) \qquad Y_{tsjdm} = \sum_{l=-1}^{3} \beta_l D_{l,jm} + \beta_x X_{tsjdm} + \theta_{sj} + \delta_{dm} + \epsilon_{tsjdm}$$

where Y_{tsjdm} is the outcome variable for transaction t in store s on date d in month m, $D_{l,jm}$ is the set of monthly event study dummies, X_{tsjdm} are control variables, θ_{sj} are store fixed effects, and δ_{dm} are date fixed effects.

Figure 12 presents the event study results, with the outcome variable being either logged transaction duration (panels a and b) or the probability of having a bagger (panel c). I juxtapose the results of using in-store data (panel a) with the results using scanner data (panel b). The scanner data comes from the full-service registers at the same three stores and on the same dates as the in-store data.[57] In both panels (a) and (b), I observe that the DCB policies led to an increase in checkout duration. Reassuringly, the $\hat{\beta}_l$ coefficients using the in-store data are comparable in size to the coefficients using scanner data.[58] These results are also consistent with the main event study results in Section 4, using the scanner data from 53 stores, in that I find a significant and persistent slowdown in transaction duration.[59]

In panel (c), using the in-store data, I find that the probability of a transaction having the assistance of a bagger temporarily decreases after the DCB policies go into effect. This could occur for several reasons. On one hand, if the same number of baggers are present after the policy as before but their presence is required for a longer period of time per transaction,

[57] With the observational data, X_{tsjdm} contains indicators for the gender and race of the person paying, whether there was a checkout interruption, and register fixed effects. With the scanner data, X_{tsjdm} contains the number of items scanned, the amount spent, and register, hour, and cashier fixed effects.

[58] In Online Appendix Table A.8, I estimate a DID model using both the scanner and in-store datasets and find similar results as these event studies.

[59] However, the $\hat{\beta}_l$ coefficients are much larger using the three store sample. In Figure 12a, $\hat{\beta}_0 = 0.173$, which is 4 times larger than what was estimated in Figure 3b, where $\hat{\beta}_0 = 0.040$. This difference may be driven by the shorter sample period of the three store data, especially in the pre-policy period (i.e., without multiple years of data and only one treated store in the sample, I am unable to fully control for seasonality and confounding factors).

they can not float to as many transactions as before the policy. Alternatively, stores may decide to use fewer baggers when their comparative advantage in packing the thin plastic bags becomes extraneous.

Finally, since I have scanner and in-store data for the same days, hours and stores, I match the scanner data transactions to their corresponding in-store data transactions. This is a challenging task as transactions that appear as one to the in-store observer may be rung up as two transactions in the scanner data, and visa versa.[60] Given the time consuming nature of matching, only in-store transactions from December 2013 (pre-policy) and January 2014 (post-policy) have been matched to the scanner data, which is roughly 41% of the transactions in the in-store sample.

From this matched data, presented in Online Appendix Table A.9, I calculate that the average plastic bag holds 3.805 items, the average paper bag holds 9.087 items, and the average brought reusable bag holds 8.744 items. Comparing transactions of similar size at the treated store, on average plastic bag transactions spend 7.244 seconds per item, paper transactions spend 8.475 seconds per item, and reusable transactions spend 7.619 seconds per item. While I cannot reject that reusable and plastic bag transactions take the same amount of time per item, I can reject at the 5% significance level that paper and plastic bag transactions take the same amount of time per item. This shows, once again, that paper bags are a slower technology, taking roughly a second more *per item* than plastic and reusable bags.

7.2 Robustness Analysis 2: Discount Chain

To explore whether this phenomenon is unique to the supermarket chain in my main analysis, I use supplementary data from a markedly different retail grocery chain. In additional to

[60] This can occur when a customer splits their purchase into smaller purchases or when a large group of customers move through the line together.

collecting observational data at the chain used throughout this paper, I collected in-store data at a discount chain within the same three treated and control California cities as describe in subsection 7.1. While the main chain is a large national chain, offering high and low prices in many products, the discount chain is a regional chain, offering name-brand products at closeout prices. Not only do these chains attract a different clientele within the same cities,[61] their management also chose different responses to the same DCB policy. The national chain chose to charge the minimum required five cents per paper bag and the discount chain chose to charge ten cents per paper bag and introduced a 15-cent thick-plastic reusable bag.

By running the same analysis on each of these chains, I am able to compare the effects of DCB polices across retail settings. I replicate the analysis in subsection 7.1 with the observational data from the discount chain.[62] Panel (d) of Figure 12 shows that DCB policies also lead to increases in transaction duration at this alternate chain. Thus the results are not exclusive to the main supermarket chain used in this paper.

7.3 Robustness Analysis 3: Washington DC Bag Tax

Are the supermarket checkout slowdowns I estimate above unique to California DCB policies, where plastic bags are banned and paper bags require a fee, or are they characteristic of other DCB policies passed in the U.S.? Furthermore, are the slowdowns unique to weekend transactions, or are they experienced in other days of the week? To answer these questions, I use scanner data from the same supermarket retailer, but for stores in the District of Columbia (DC) metropolitan area. While California has favored using plastic bag bans and paper bag fees because a state law temporarily prohibited the taxing of plastic bags,[16] local

[61] The discount chain has a 15 percentage point greater share of racial minority customers than the national chain.

[62] Please see Online Appendix E for a more detailed description of the variables in the in-store data at the discount chain and Appendix Table A.8 for the results of a DID analysis using these discount chain data.

governments in other states have had more flexibility in their policy tool options.[63] On January 1, 2010, DC enacted a bag tax, requiring all stores that sell food items to charge a 5-cent tax per plastic or paper bag issued.[64] While bag bans are command-and-control policies that regulate behavior directly by changing the permissible choice set, bag taxes are market-based policies that leave the choice set intact but change prices in order to incentivize consumers to alter their own behavior.

Using scanner data from DC, I examine the effects of a different type of DCB policy in a different region, in order to learn about the generalizability of the results. The DC scanner dataset covers 4 months in the pre-tax period (Dec. 2008, Jan. 2009, Feb. 2009, Dec. 2009) and 5 months in the post-tax period (Jan. 2010, Feb. 2010 Dec. 2010, Jan. 2011 and Feb. 2011). The sample includes transactions during the peak hours of 3:00-7:00pm for every day of the week during these months. Having scanner data from all days of the week allows me to verify whether the policy effects are similar between weekend and weekday transactions. This partially addresses a concern about the external validity of the main results, where the scanner data sample was restricted to weekend transactions. Finally, the sample includes stores within a 30 mile radius of DC that are open without interruption between December 2008 and January 2011. This gives me 8 treated stores in DC and 31 control stores in Virginia and Maryland.[65]

---

[63] Conversely, some states (including Arizona, Idaho, Michigan, and Missouri) have passed laws that ban local governments from banning or taxing plastic bags ("State Plastic and Paper Bag Legistlation; Fees, Taxes and Bans | Recycling and Reuse." *National Conference of State Legislatures.* Nov. 11, 2016. Online, *accessed Dec. 18, 2016*).

[64] Previous studies have found that the vast majority of customers choose plastic bags over paper bags when given the option (Homonoff, 2016; Taylor and Villas-Boas, 2016b).

[65] Online Appendix Table A.10 presents the average characteristics for treated and control stores in the pre-policy period. While treated and control stores are balanced with respect to building characteristics, treated stores are located in census block-groups with lower median incomes, higher shares of black residents, and lower rates of vehicle ownership. With respect to transaction-level characteristics, average pre-policy transactions at treated stores take slightly longer to complete, yet have roughly the same size and cost, as at control stores. I drop transactions occurring at self-checkout registers because only 18 percent of stores have self-checkout during my sample period. I also drop transactions occurring on Christmas (Dec. 25) and

I estimate the effect of the DC bag tax on transaction duration with data averaged to the store-week level and the following event study model:

$$Y_{sjw} = \sum_{l=-5}^{8} \beta_l D_{l,jw} + \beta_x X_{sjw} + \theta_{sj} + \chi_w + \epsilon_{sjw} \tag{7}$$

where Y_{sjw} is the logged transaction duration in store s, jurisdiction j, and week-of-sample w, $D_{l,jw}$ are indicators for transactions at the treated stores in DC during the weeks before and after the bag tax, X_{sjw} is the set of control variables, θ_{sj} are store fixed effects, and δ_w are week-of-sample fixed effects. $D_{-5,jw}$ equals one for all weeks between December 2008 and February 2009 (i.e., a year before policy implementation). Similarly, $D_{8,jw}$ equals one for all weeks between December 2010 and February 2011 (i.e., a year after policy implementation). The week prior to implementation ($l = -1$) is the omitted category.

Figure 13 plots the results using the DC scanner data. Panel (a) shows that, similar to Californian bag bans, the DC bag tax also led to slower transactions. The slowdown peaks in the second week of the policy, with transactions taking 8.6% longer to complete. However, unlike the California bag bans shown in Figure 3, the slowdown from the DC bag tax lessens over the sample period, with $\hat{\beta}_8 = 0.036$ (a year after the policy) half the magnitude of $\hat{\beta}_1 = 0.086$.[66] While the slowdown from the DC bag tax diminishes over time, transactions in DC stores remain persistently (and statistically significantly) slower than control stores a year after the policy. Thus both bag bans and bag fees are found to cause persistent slowdowns in checkout duration.[67]

during the 2009 and 2010 North American blizzards (Dec. 19–20, 2009; Feb. 6-7, 10-11, 2010), as many of the stores were closed on these dates.

[66] I can reject that $\hat{\beta}_1 = \hat{\beta}_8$ at the 5% significance level.

[67] In Online Appendix Figure A.3, I also replicate the analyses in Figures 5(a), 6(a), and 6(b) using the DC scanner data. This appendix figure shows: (a) 36% of transactions in DC pay for a disposable bag a year after the policy, (b) there is no statistically significant slowdown in transaction duration for the transactions that do not pay the bag tax, and conversely, (c) the average transaction that pays the bag tax is 13.0% slower in the first week of the policy and 7.9% slower a year after the policy.

Next I examine heterogeneity by day-of-the-week. I estimate a difference-in-differences form of Equation 7, where the event dummies are collapsed to a single indicator, D_{jw}, equal to 1 if jurisdiction j has a bag tax effective in week-of-sample w. Figure 13(b) presents the β_D coefficients, estimated using the full sample and separately by day-of-the-week. The outcome variable is once again logged transaction duration. Using the full sample of data, the DC bag tax led to an average 3.8% increase in transaction duration. Separating the sample by day-of-the-week reveals little to no difference in policy effects between weekday and weekend transactions. The coefficients by day-of-the-week, while ranging from 3.0% on Sundays to 4.9% on Tuesdays, are statistically indistinguishable from each other.[68] Consequently, this robustness analysis provides strong supportive evidence that the main results—using scanner data from weekends only—are externally valid to other days of the week.

8 Conclusion

This study is the first to quantify a hidden time cost of a popular environmental policy aimed at altering consumer behavior. Using detailed scanner data and an event study design, I find that DCB policies lead to a persistent 3% increase in supermarket checkout duration. While I observe evidence of learning at the cashier level, this learning does not offset the slowdown from the policies, which persists over the entire sample period. The time costs are heterogeneous by whether paper bags are purchased, with the transactions paying the paper bag fee experiencing a 9.2% increase in transaction duration. This is due to paper bags being slower to pack per item than both disposable plastic bags and reusable bags.

Second, I find that DCB policies lead to significant decreases in sales, with supermarkets and grocery stores in treated jurisdictions experiencing a persistent 1.6% drop in sales.

[68] Variation in policy effects by day-of-the-week reflects the share of customers paying the plastic bag tax each day, which is lowest on Sundays (36% of customers) and highest on Tuesdays (41% of customers).

Aggregating to the state level, this would translate to a $2.3 billion loss in sales per year for Californian grocers. Additional analyses suggest these food dollars are shifting from grocers to restaurants and other food-away-from-home establishments. Using a hazard model, I show that 10% of the reduction in sales from DCB policies is driven by the 3% slowdown in transaction duration. Thus not only do DCB policies impose a time cost on customers, customers respond to this time cost by shopping less frequently at affected stores. Future work should analyze mechanisms behind the remaining 90% of lost sales from the policies, which could be driven by the emotional costs of DCB policies (e.g., reactance, fear of being stigmatized) or the time and effort costs of DCB policies in aspects of grocery shopping other than checkout.

In evaluating DCB policies, their success has been measured primarily as the reduction in plastic bags consumed. Based on the in-store observational data of bag use at checkout, I calculate that a statewide DCB policy would lead to 5.3 billion fewer disposable bags used per year in California.[69] Using an estimate of how much taxpayers spend in collection, processing, and landfilling disposable bag waste, this plastic bag reduction would save $58.3 million in tax dollars annually.[70] In comparison, a 3% slowdown of all grocery trips per year would cost Californian shoppers 1.42 million hours per year, or $18.9 million at half the average hourly wage. Therefore, not including the checkout time cost of DCB policies overstates the net benefits of the policies by 32%. Granted, these estimates do not include the environmental benefits of reduced plastic marine debris, which has been greatly espoused but not yet quantified, nor does it include the potential costs of increased away-from-home meal consumption. Thus, for these policies to be welfare enhancing, the long run environmental

[69] In the in-store data, I find that the average customer uses 3.73 fewer disposable bags per transaction post-DCB policy. Given Californian adults make 1.42 billion grocery transactions per year, this equal 5.3 billion fewer disposable bags per year.

[70] Herrera Environmental Consultants, Inc. (2008) find that the collection, processing, and landfilling of disposable bag waste costs local governments 1.1 cents per bag.

benefits of less plastic in oceans and waterways needs to be larger than the time costs discussed in this paper, as well as the potential health impacts of fewer grocery sales.

For policymakers concerned with both consumer waste and public health, this paper reveals that incentivizing environmentally-friendly behaviors can be at odds with making healthy eating more convenient. Even though disposable carryout bags comprise a very small fraction of the monetary expenditures for food production, their role in reducing the time necessary for healthy food acquisition is not trivial, especially when aggregated across all consumers and all consumption occasions. Economic incentives and regulations which seem like highly-effective, low-cost behavioral nudges, may also have large non-monetary costs, and ignoring these costs overstates the welfare gains of such policies. While this paper quantifies a non-monetary cost of an environmental policy, I do not complete a full welfare analysis, nor do I entirely explore all the ways consumers react to this policy. Future work should endeavor to quantify other non-monetary costs of behavioral change (e.g., emotional responses, mental effort costs) as important mechanisms behind policy circumvention, and examine heterogenity in consumer responses to non-monetary costs.

References

Adda, J. and F. Cornaglia (2010). The Effect of Bans and Taxes on Passive Smoking. *American Economic Journal: Applied Economics 2*(1), 1–32.

Allcott, H. (2016). Paternalism and Energy Efficiency: An Overview. *Annual Review of Economics 8*, 145–176.

Allcott, H. and J. B. Kessler (2018). The Welfare Effects of Nudges: A Case Study of Energy Use Social Comparisions. *American Economic Journal: Applied Economics, forthcoming*.

Arrow, K. J. (1962). The Economic Implications of Learning by Doing. *Review of Economic Studies 29*(3), 155–173.

Ashenmiller, B. (2011). The Effect of Bottle Laws on Income: New Empirical Results. *American Economic Review: Papers & Proceeding 101*(3), 60–64.

Beatty, T., P. Berck, and J. Shimshack (2007). Curbside Recycling in the Presence of Alternatives. *Economic Inquiry 45*(4), 739–755.

Becker, G. (1965). A Theory of the Allocation of Time. *Economic Journal 75*(299), 493–517.

Becker, G. S., M. Grossman, and K. M. Murphy (1991). Rational Addiction and the Effect of Price on Consumption. *The American Economic Review: Papers and Proceedings 81*(2), 237–241.

Bertrand, M., E. Duflo, and S. Mullainathan (2004). How Much Should We Trust Difference-in-Differences Estimates? *Quarterly Journal of Economics 119*(1), 249–275.

Burnett, H. S. (2013). Do Bans on Plastic Grocery Bags Save Cities Money? Policy Report No. 353, National Center for Policy Analysis.

California Senate Rules Committee (2014). Solid Waste: Single-use Carryout Bags. Senate Bill 270, Bill Analysis.

Chaloupka, F. J. (1991). Rational Addictive Behavior and Cigarette Smoking. *Journal of Political Economy 99*(4), 722–742.

Choi, J. J., D. Laibson, B. Madrian, and A. Metrick (2003). Optimal Defaults. *American Economic Review 93*(2), 180–185.

Clapp, J. and L. Swanston (2009). Doing Away with Plastic Shopping Bags: International Patterns of Norm Emergence and Policy Implementation. *Environmental Politics 18*(3), 314–332.

Conley, T. (2008). *Spatial Econometrics*. Houndsmills: Palgrave Macmillan.

Convery, F., S. McDonnell, and S. Ferreira (2007). The Most Popular Tax in Europe? Lessons from the Irish Plastic Bag Levy. *Environmental and Resource Economics 38*(1), 1–11.

Correia, S. (2014). REGHDFE: Stata Module to Perform Linear or Instrumental-Variable Regression Absorbing Any Number of High-Dimensional Fixed Effects. Statistical Software Components s457874, Boston College Department of Economics, revised 25 Jul. 2015.

Craig, A., E. Garbarino, S. A. Heger, and R. Slonim (2016). Waiting to Give: Stated and Revealed Preferences. *Management Science 63*(11), 3672–3690.

Damgaard, M. T. and C. Gravert (2016). The Hidden Costs of Nudging: Experimental Evidence from Reminders in Fundraising. Available at SSRN: https://ssrn.com/abstract=2747893.

Davis, L. (2008). The Effect of Driving Restrictions on Air Quality in Mexico City. *Journal of Political Economy 116*(1), 38–81.

Dikgang, J., A. Leiman, and M. Visser (2012). Elasticity of Demand, Price and Time: Lessons from South Africa's Plastic-bag Levy. *Applied Economics 44*(26), 3339–3342.

Dupas, P. (2014). Short-run Subsidies and Long-run Adoption of New Health Products: Evidence from a Field Experiment. *Econometrica 82*(1), 197–228.

Fetzer, T. (2014). Can Workfare Programs Moderate Violence? Evidence from India. STICERD Working Paper.

Fowlie, M., M. Greenstone, and C. Wolfram (2015). Are the Non-Monetary Costs of Energy Efficiency Investments Large? Understanding Low Take-up of a Free Energy Efficiency Program. *American Economic Review: Papers & Proceedings 105*(5), 201–204.

Freinkel, S. (2011). *Plastic: A Toxic Love Story.* Houghton Mifflin Harcourt.

Fullerton, D. and T. C. Kinnaman (1996). Household Responses to Pricing Garbage by the Bag. *American Economic Review 86*(4), 971–984.

Gicheva, D., J. Hastings, and S. Villas-Boas (2010). Investigating Income Effects in Scanner Data: Do Gasoline Prices Affect Grocery Purchases? *American Economic Review 100*(2), 480–484.

Goodman, J. (2008). Who Does the Grocery Shopping, and When Do They Do It? White paper, The Time Use Institute. http://timeuseinstitute.org/GroceryWhitePaper2008.pdf, accessed Mar. 21, 2016.

Hamrick, K., M. Andrews, J. Guthrie, D. Hopkins, and K. McClelland (2011). How Much Time Do Americans Spend on Food. Technical report, EIB-86, United States Department of Agriculture, Economic Research Service.

Herrera Environmental Consultants, Inc. (2008). Alternatives to Disposable Shopping Bags and Food Service Items, Volume I and II. Prepared for Seattle Public Utilities.

Homonoff, T. (2016). Can Small Incentives Have Large Effects? The Impact of Taxes Versus Bonuses on the Demand for Disposable Bags. Working Paper.

Hsiang, S. M. (2010). Temperatures and Cyclones Strongly Associated with Economic Production in the Caribbean and Central America. *Proceddings of the National Academy of Sciences of the United States of America 107*(35), 15367–15372.

Jambeck, J. R., R. Geyer, C. Wilcox, T. R. Siegler, M. Perryman, A. Andrady, R. Narayan, and K. L. Law (2015). Plastic Waste Inputs from Land into the Ocean. *Science 347*(6223), 768–771.

Johnson, E. and D. Goldstein (2003). Do Defaults Save Lives. *Science 302*(5649), 1338–1339.

Just, D. R. and A. S. Hanks (2015). The Hidden Cost of Regulation: Emotional Responses to Command and Control. *American Journal of Agricultural Economics 97*(5), 1385–1399.

Kahneman, D. (2011). *Thinking, Fast and Slow.* Farrar, Straus and Giroux.

Kant, A. K., M. I. Whitley, and B. I. Graubard (2015). Away from Home Meals: Associations with Biomarkers of Chronic Disease and Dietary Intake in American Adults, NHANES 20052010. *International Journal of Obesity 39*, 820–827.

Katz, K. L., B. M. Larson, and R. C. Larson (1991). Prescription for the Waiting-in-Line Blues: Entertain, Enlighten, and Engage. *MIT Sloan Management Review 32*(2), 44–53.

Krueger, A. (2009). A Hidden Cost of Health Care: Patient Time. *New York Times* (Feb. 9), https://economix.blogs.nytimes.com/2009/02/09/a-hidden-cost-of-health-care-patient-time/, accessed Dec. 14, 2017.

Larson, R. (1987). Perspectives on Queues: Social Justice and the Psychology of Queuing. *Operations Research 35*(6), 895–905.

Madrian, B. and D. F. Shea (2001). The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior. *Quarterly Journal of Economics 116*(4), 1149–1187.

Maister, D. H. (1985). *The Psychology of Waiting in Line*. Lexington Books: Lexington, MA.

Mathiowetz, N. and G. Duncan (1988). Out of Work, Out of Mind: Response Errors in Retrospective Reports of Unemployment. *Journal of Business and Economic Statistics 6*(2), 221–229.

Neter, J. and J. Waksberg (1964). A Study of Response Errors in Expenditures Data from Household Interviews. *Journal of the American Statistical Association 59*(305), 18–55.

Rosenberg, N. (1982). *Inside the Black Box: Technology and Economics*. Cambridge University Press, New York, NY.

Small, K. (1992). *Urban Transportation Economics*. Harwood Academic Publishers: Chur, Switzerland.

Small, K. and E. Verhoef (2007). *The Economics of Urban Transportation*. New York: Routledge.

Taylor, R. and S. Villas-Boas (2016a). Food Store Choices of Poor Households: A Discrete Choice Analysis of the National Household Food Acquisition and Purchase Survey (FoodAPS). *American Journal of Agricultural Economics 98*(2), 513–532.

Taylor, R. and S. B. Villas-Boas (2016b). Bans vs. Fees: Disposable Carryout Bag Policies and Bag Usage. *Applied Economic Perspectives and Policy 38*(2), 351–372.

Tom, G. and S. Lucey (1995). Waiting Time Delays and Customer Satisfaction in Supermarkets. *Journal of Service Marketing 9*(5), 20–29.

Van Riel, A. C., J. Semeijn, D. Ribbink, and Y. Bomert-Peters (2012). Waiting for Service at the Checkout: Negative Emotional Responses, Store Image and Overall Satisfaction. *Journal of Service Management 23*(2), 144–169.

Wilcox, C., N. J. Mallos, G. H. Leonard, and A. Rodriguez (2016). Using Expert Elicitation to Estimate the Impacts of Plastic Pollution on Marine Wildlife. *Marine Policy 65*, 107–114.

Yaktine, A. and J. Caswell (2013). *Supplemental Nutrition Assistance Program: Examining the Evidence to Define Benefit Adequacy*. National Academies Press, Washington, DC.

Table 1: Average Store Characteristics and Demographics

	(1) Control Stores	(2) Treat Stores	(3) P-value of Diff.	(4) Calif.	(5) U.S.
Store Characteristics					
Building Size (ft^2)	41,782.75	43,434.45	0.632		
Open Year	1985	1984	0.778		
Last Remodel Year	2005	2005	0.992		
Departments & Services (share)					
Bakery	0.75	0.79	0.755		
Pharmacy	0.50	0.64	0.338		
Deli	1.00	0.94	0.270		
Floral	0.90	0.91	0.915		
Coffee Bar	0.70	0.52	0.193		
Gas Station	0.05	0.06	0.874		
Juice Bar	0.10	0.09	0.915		
Sandwich Counter	0.05	0.06	0.874		
Self-checkout registers (share)	0.45	0.48	0.810		
Store Location Demographics					
Median Income ($)	$63,120	$62,280	0.874	$47,493	$41,994
Household Size (#)	2.61	2.58	0.726	2.87	2.59
White (share)	0.70	0.69	0.785	0.60	0.75
Black (share)	0.06	0.05	0.723	0.07	0.12
Asian (share)	0.11	0.12	0.590	0.11	0.04
Over 65 (share)	0.12	0.12	0.956	0.11	0.12
Do not own vehicle (share)	0.06	0.06	0.415	0.10	0.10
Urban (share)	0.75	0.82	0.562	0.87	0.79
N Stores	20	33			

Source: Store characteristic data were provided by the retailer. Store demographic data come from Gicheva et al. (2010), who use 2000 US Census data for each store's census block-group. The state and country level data come from the 2000 US Census, *Table DP-1. Profile of General Demographic Characteristics: 2000*; Geographic Areas: California and United States. Urban areas are locations with populations densities greater than 500 people per square mile.

Table 2: Average Transaction and Store-Shift Characteristics in 2011

Transaction-Level Characteristics	Control	Treat	P-value of Diff.
Transaction Duration (minutes)	**1.97**	**2.08**	**0.445**
Full-Service	1.89	1.99	0.160
Express	1.38	1.48	0.033
Self-Checkout	3.70	3.98	0.078
Items Scanned (#)	**13.17**	**13.44**	**0.629**
Full-Service	18.65	19.22	0.658
Express	7.64	8.57	0.046
Self-Checkout	6.45	6.62	0.473
Amount Paid ($)	**38.83**	**39.69**	**0.663**
Full-Service	54.60	56.38	0.662
Express	23.35	25.89	0.093
Self-Checkout	19.64	19.87	0.750
Store-Shift Characteristics			
Transactions per shift (#)	542.98	566.78	0.620
Amount spent per shift ($)	20,601.71	21,811.17	0.641
Items bought per shift (#)	6,952.38	7,380.48	0.614
Registers open per shift (#)	7.89	8.54	0.515
Register Capacity (#)	10.25	10.64	0.700
N Stores	20	33	–
N Stores w/Self-checkout	9	16	–

Note: 2011 is in the pre-policy period for all stores in the sample. *Source*: Author's calculations from the scanner data.

Table 3: Effect of DCB Policies on Transactions per Shift and Line Length
(*Store-Week Averages*)

	(1) Txn. Duration (min.)	(2) Txns. per Shift (#)
Levels (Y_{sjw})		
DCB Policy Effective (=1)	0.064	-18.226
	(0.017)	(6.047)
Percent (lnY_{sjw})		
DCB Policy Effective (=1)	0.033	-0.030
	(0.008)	(0.010)
N Obs.	9381	9381
Standard Errors	Cluster	Cluster
Covariates X_{sjw}	Yes	Yes
Store FE	Yes	Yes
Week-of-sample FE	Yes	Yes
Mean Y_{sjw} (2011)	2.045	556.829

Note: Coefficients from DID Equation 2, with the outcome variable estimated in both levels and logs. In column (1), the outcome variable is the average transaction duration, measured in minutes, in store s, jurisdiction j, and week-of-sample w. In column (2), the outcome variable is the average number of transactions completed per 3 hour shift in store s, jurisdiction j, and week-of-sample w. Control variables, X_{sjw}, include average transaction size, average transaction expenditures, and the share of transactions purchasing each of the following items—produce, meat/seafood, dairy/refrigerated, frozen, bakery/deli, shelf-stable food, alcohol/tobacco, infant/toddler, floral department, and pet items. Standard errors are in parentheses. Standard errors are estimated using two-way error clustering at the policy jurisdiction and week-of-sample level.

Table 4: Average Effect of DCB Policies and Transaction Duration on Likelihood to Return

Likelihood to Return	(1)	(2)	(3)	(4)
DCB Policy Effective (=1)	-0.009	-0.010		-0.009
	(0.005)	(0.005)		(0.005)
Txn. Duration (min.)	-0.011		-0.011	-0.011
	(0.002)		(0.002)	(0.002)
Ban Eff. X Txn. Duration				0.002
				(0.003)
N Obs.	8109654	8109654	8109654	8109654
Ancillary parameter ($\hat{\gamma}$)	-0.003	-0.003	-0.003	-0.003
Log Psuedolikihood	-600499.483	-601104.134	-600522.044	-600495.165
Standard Errors	Cluster	Cluster	Cluster	Cluster
Store FE	Yes	Yes	Yes	Yes
Week-of-Sample FE	Yes	Yes	Yes	Yes

Note: Coefficients of survival model (Equation 5) with Gompertz parametrization. The outcome variable is the likelihood of customer card user i to return to store s in jurisdiction j on week-of-sample w, given it is t weeks since their last visit. Standard errors in parentheses are clustered at the jurisdiction level.

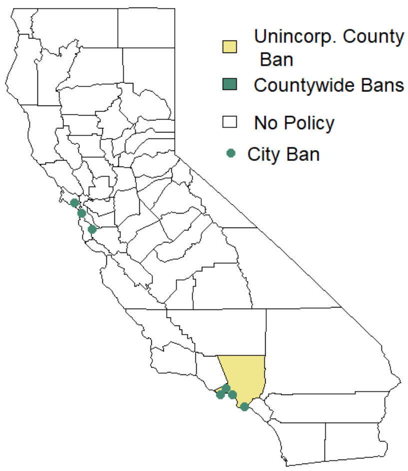
Table 5: Average Transaction Duration (*In-store vs. Scanner Data*)

	In-store Data	Scanner Data	Difference
Full Sample Mean	1.718	1.837	-0.119
(SD)/[SE]	(1.114)	(1.234)	[0.031]
N Obs.	1,692	34,028	
Stores with DCB Policy	1.756	1.910	-0.154
(SD)/[SE]	(1.144)	(1.252)	[0.042]
N Obs.	934	17,562	
Stores without DCB Policy	1.670	1.759	-0.089
(SD)/[SE]	(1.073)	(1.210)	[0.045]
N Obs.	758	16,466	

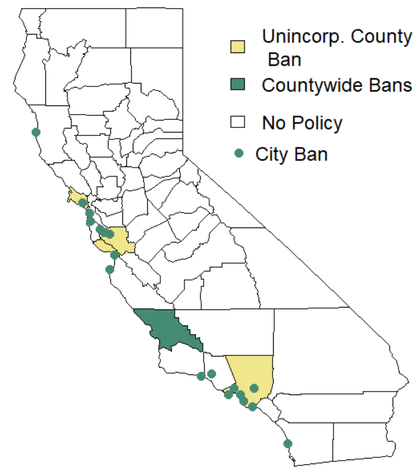
Note: Standard deviations in parentheses. Standard errors of difference in brackets. *Source*: Author's calculations from observational data collected in-store and from the scanner data corresponding to the same days and stores as the observational data.

Figure 1: California Disposable Carryout Bag (DCB) Policies over Time

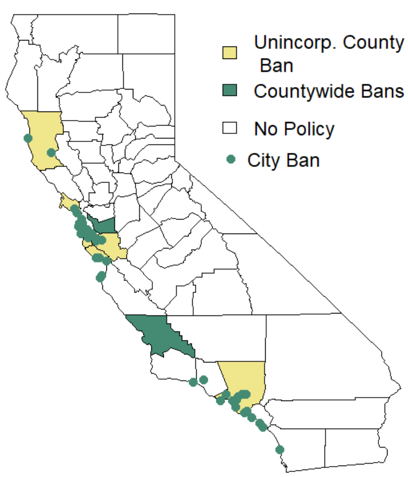
(a) Policies Implemented Before 2012



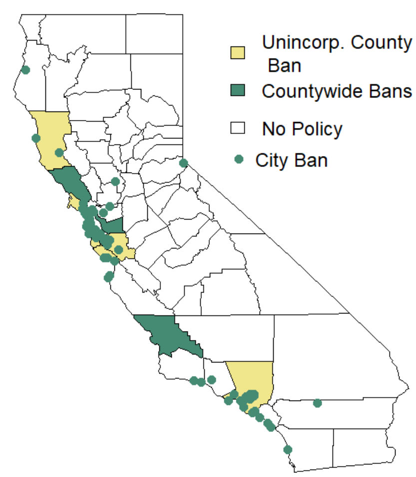
(b) Policies Implemented Before 2013



(c) Policies Implemented Before 2014

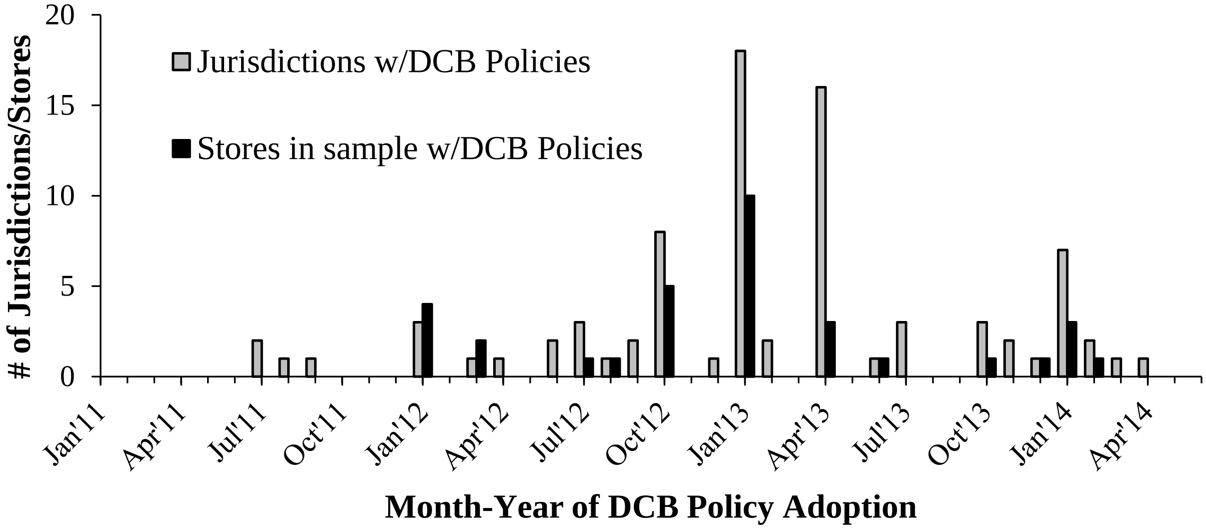


(d) Policies Implemented Before 2015



Note: The local governments of unincorporated counties and incorporated cities can pass ordinances to regulate disposable carryout bags. Countywide policies occur when all cities and unincorporated areas in a county pass DCB regulations.

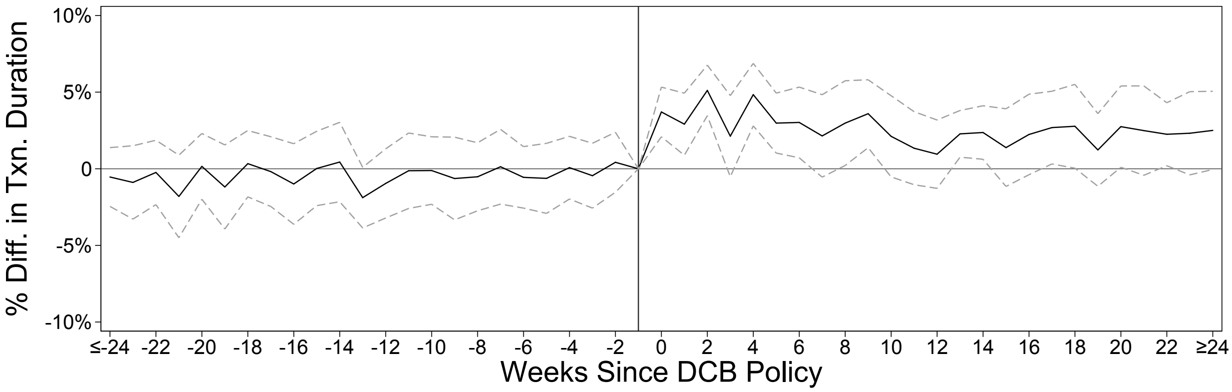
Figure 2: Number of California Jurisdictions Implementing DCB Policies by Month and Year



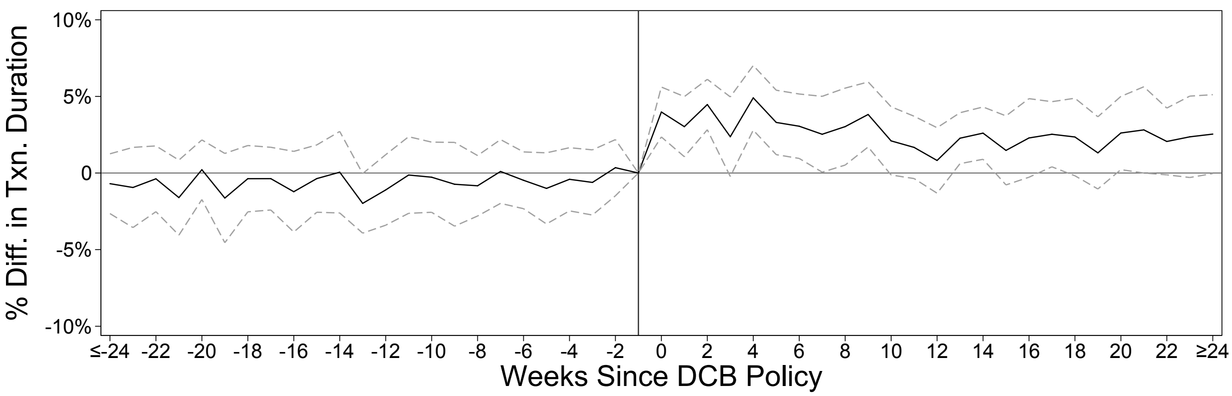
Source: Author's calculations.

Figure 3: Effect of DCB Policies on Transaction Duration (*Store-Week Averages*)

(a) Logged Transaction Duration—Without Control Variables

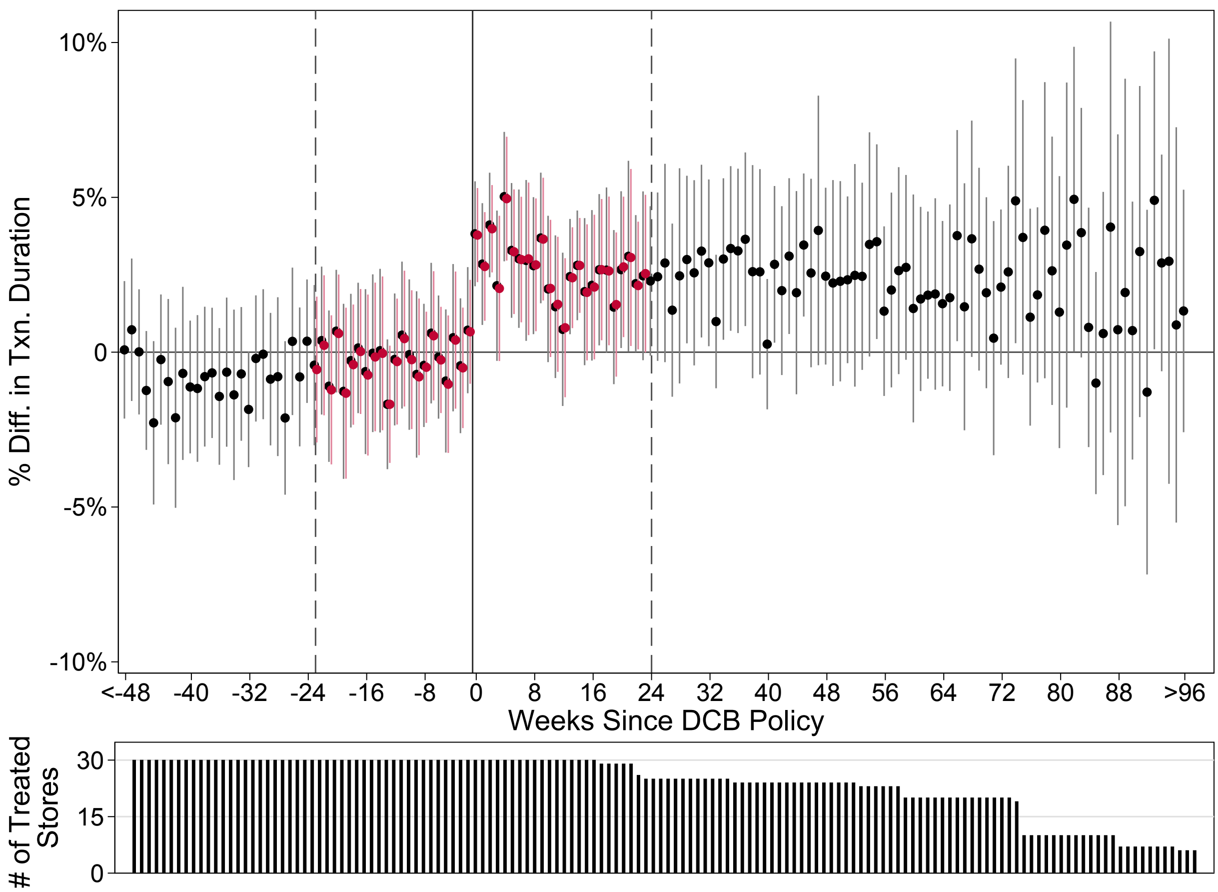


(b) Logged Transaction Duration—With Control Variables



Note: The figure panels display the $\hat{\beta}_l$ coefficient estimates from event study Equation 1. The dependent variable is logged average transaction duration, measured in minutes, in store s, jurisdiction j, and week-of-sample w. Upper and lower 95% confidence intervals are depicted in gray, estimated using two-way cluster robust standard errors on policy jurisdiction and week-of-sample. Panel (a) presents the specification of Equation 1 with event study indicators, store fixed effects, and week-of-sample fixed effects. The specification in panel (b) additionally includes control variables, X_{sjw}, for average transaction size, average transaction expenditures, and the share of transactions purchasing each of the following items—produce, meat/seafood, dairy/refrigerated, frozen, bakery/deli, shelf-stable food, alcohol/tobacco, infant/toddler, floral department, and pet items.

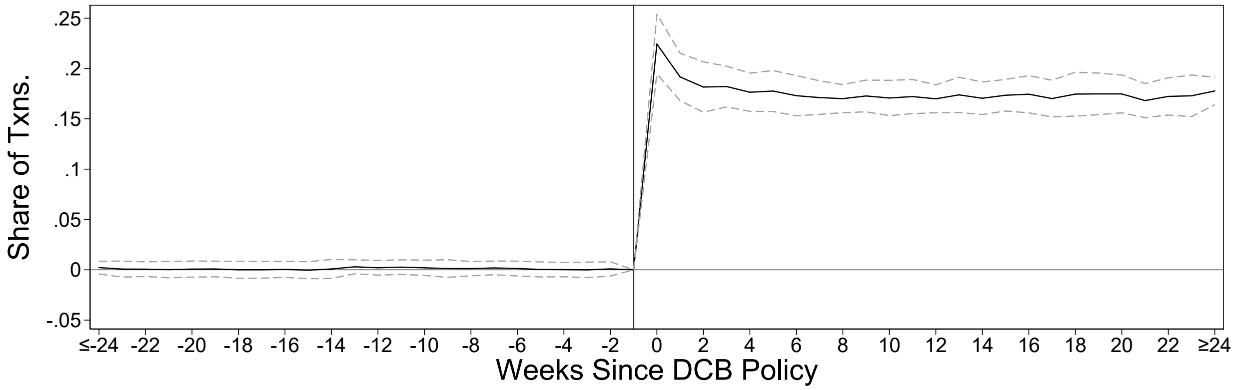
Figure 4: Effect of DCB Policies on Transaction Duration - Extended Endpoints (*Store-Week Averages*)



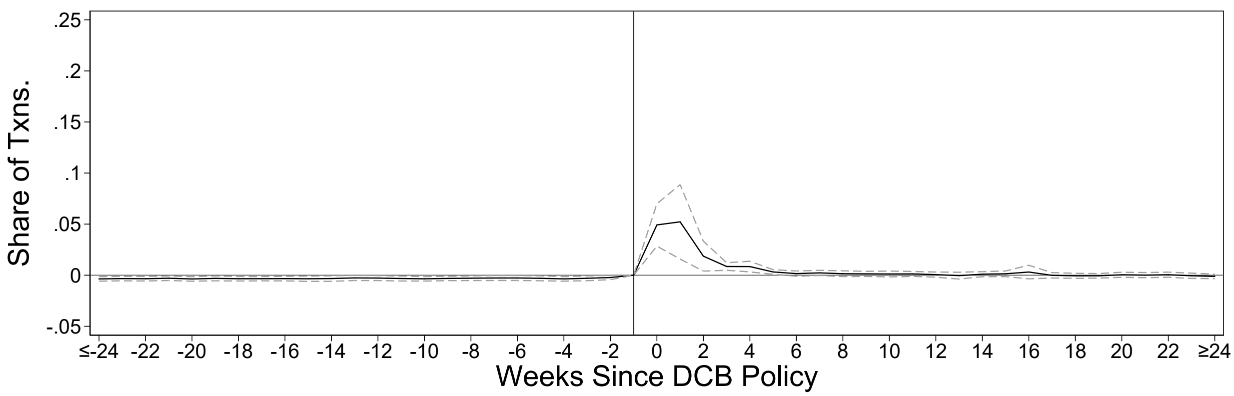
Note: The top panel of the figure displays the $\hat{\beta}_l$ event study estimates, in black, from the extended event study equation: $Y_{sjw} = \sum_{l=-49}^{97} \beta_l D_{l,jw} + \beta_x X_{sjw} + \theta_{sj} + \delta_w + \epsilon_{sjw}$. In red, between the dashed lines, are the β_l estimates from Equation 1—same as in Figure 3a—where the event study endpoints are instead binned at ± 24 weeks. The dependent variable is logged average transaction duration, measured in minutes, in store s, jurisdiction j, and week-of-sample w. Upper and lower 95% confidence intervals are depicted in gray, estimated using two-way cluster robust standard errors on policy jurisdiction and week-of-sample. The bottom panel presents a bar plot for the number of treated stores with $D_{l,sw} = 1$. Stores that implement policies later in the sample period mechanically have fewer post-policy weeks than stores with early implementation dates. This plot shows that while all 33 treated stores have $D_{15,sw} = 1$, only 7 treated stores have $D_{96,sw} = 1$.

Figure 5: Effect of DCB Policies on Share of Transactions Purchasing Paper Bags and Reusable Bags (*Store-Week Averages*)

(a) Share Purchasing Paper Bags



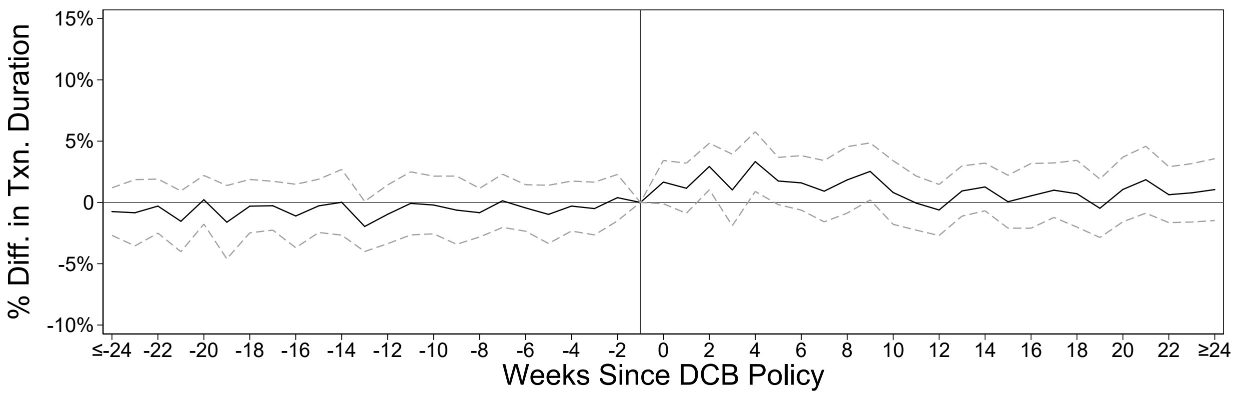
(b) Share Purchasing Reusable Bags



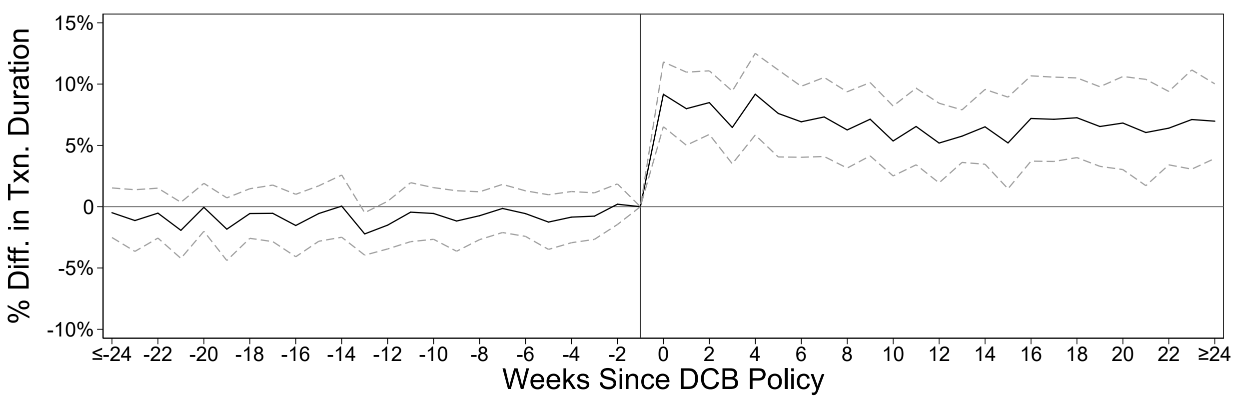
Note: The figure panels display the $\hat{\beta}_1$ coefficient estimates from the full specification of event study Equation 1. The dependent variable is share of transactions in store s, jurisdiction j, and week-of-sample w either purchasing paper bags (panel a) or purchasing reusable bags (panel b). Upper and lower 95% confidence intervals are depicted in gray, estimated using two-way cluster robust standard errors on policy jurisdiction and week-of-sample.

Figure 6: Heterogeneity in Effect of DCB Policies on Transaction Duration, by Paper Bag Purchase (*Store-Week Averages*)

(a) Without paying paper bag fee



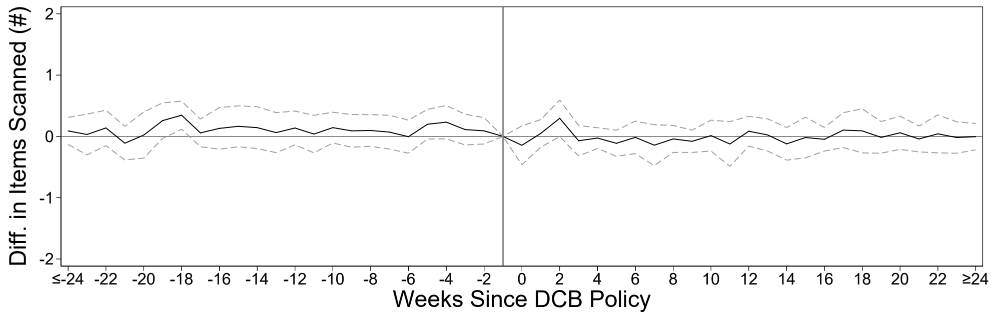
(b) Paid paper bag fee



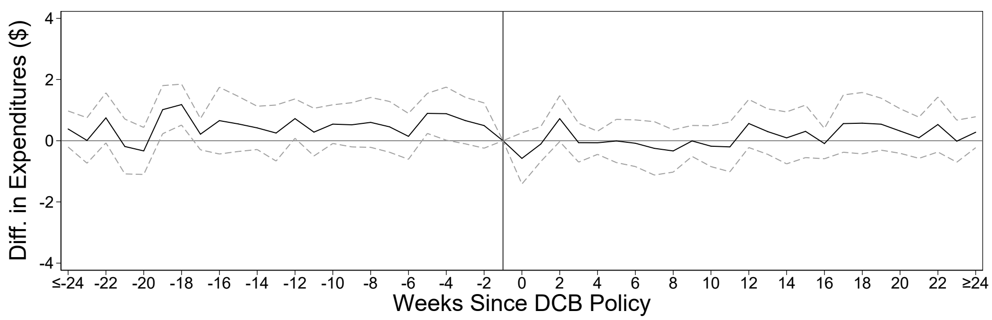
Note: The figure panels display the $\hat{\beta}_1$ coefficient estimates from the full specification of event study Equation 1. The dependent variable is logged average transaction duration, measured in minutes, in store s, jurisdiction j, and week-of-sample w, by paper bag use. Upper and lower 95% confidence intervals are depicted in gray, estimated using two-way cluster robust standard errors on policy jurisdiction and week-of-sample.

Figure 7: Effect of DCB Policies on Number of Items Scanned and Amount Spent (*Store-Week Averages*)

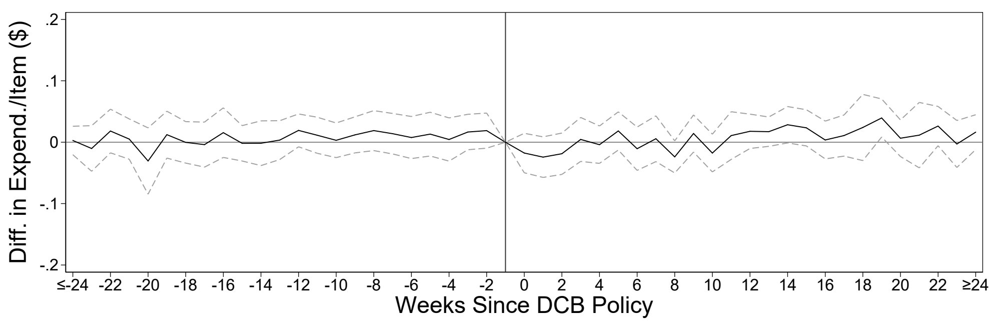
(a) Items Scanned per Transaction, Not Including Carryout Bags (#)



(b) Amount Spent per Transaction ($)



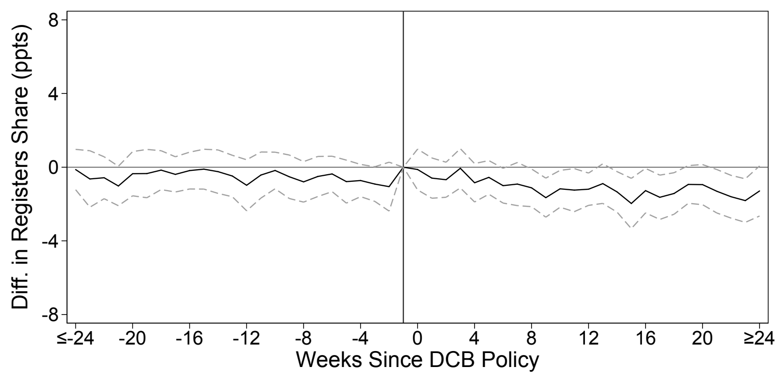
(c) Amount Spent per Item ($)



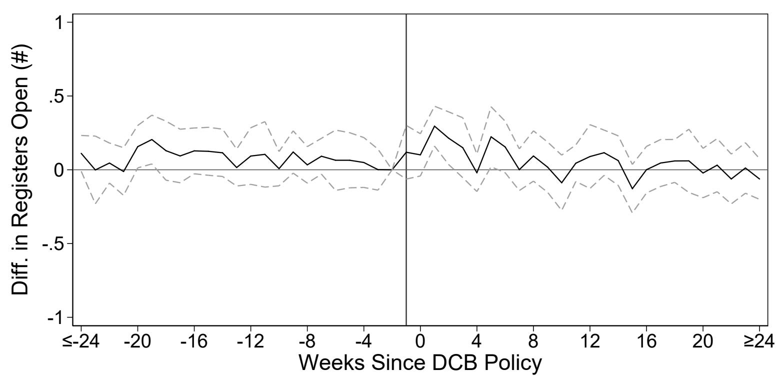
Note: The figures display the $\hat{\beta}_l$ coefficient estimates from the simple specification of event study Equation 1, controlling for only store and week-of-sample fixed effects. The dependent variables for store s, jurisdiction j, and week-of-sample w are (a) the average number of items scanned per transaction, not including carryout bags, (b) the average amount spent per transaction, and (c) the average amount spent per item. Upper and lower 95% confidence intervals are depicted in gray, estimated using two-way cluster robust standard errors on policy jurisdiction and week-of-sample.

Figure 8: Effect of DCB Policies on Register Choice and Registers Open (*Store-Week Averages*)

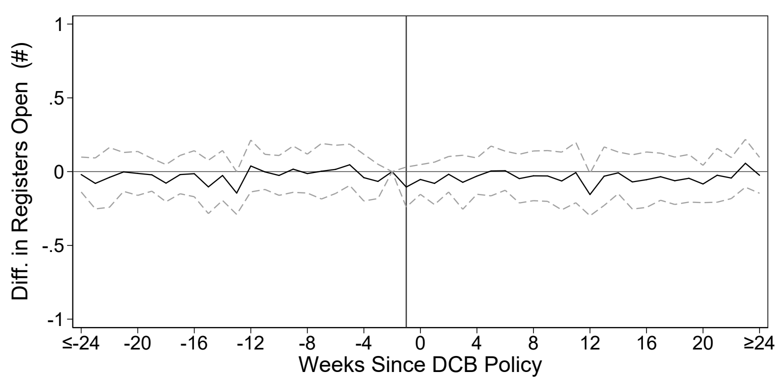
(a) Share of Transactions at Cashier-operated Registers versus Self-checkout Registers



(b) Number of Cashier-operated Registers Open

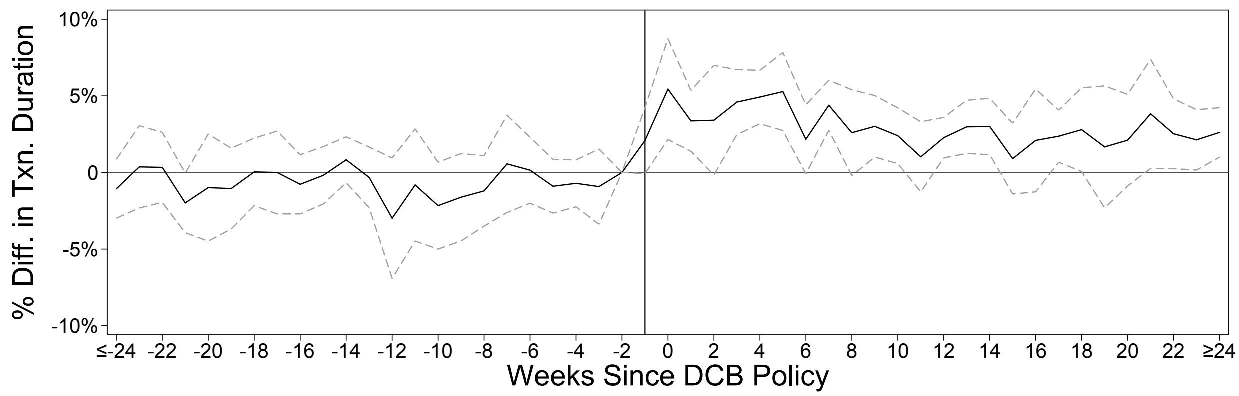


(c) Number of Self-checkout Registers Open



Note: The figures display the $\hat{\beta}_l$ coefficient estimates from the simplest specification of Equation 1. The dependent variables in panel (a) is the share of transactions at cashier-operated registers (versus self-checkout registers) in store s, jurisdiction j, and week-of-sample w. The dependent variables in panels (b) and (c) are the average number of cashier-operated and self-checkout registers open in store s, jurisdiction j, and week-of-sample w. Upper and lower 95% confidence intervals are depicted in gray, estimated using two-way cluster robust standard errors on policy jurisdiction and week-of-sample.

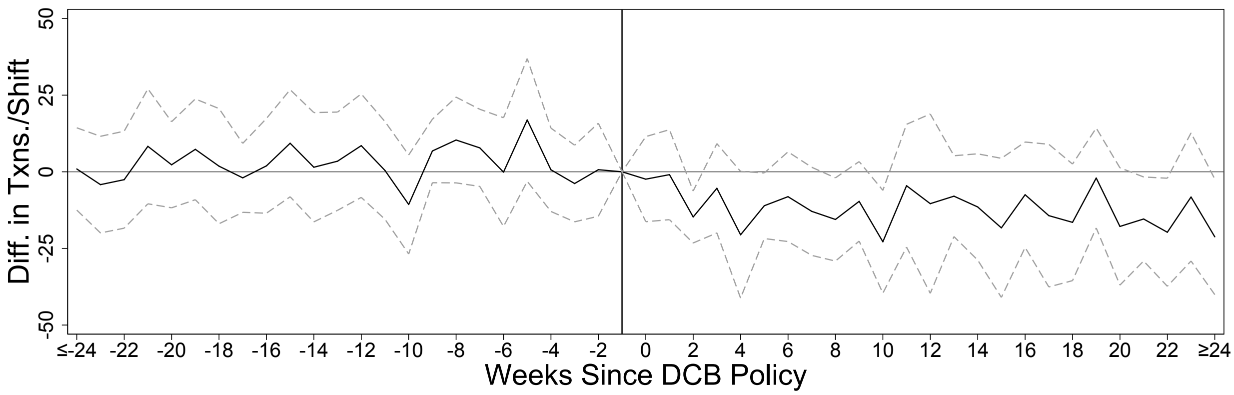
Figure 9: Effect of DCB Policies on Transaction Duration (*Cashier-Week Averages*)



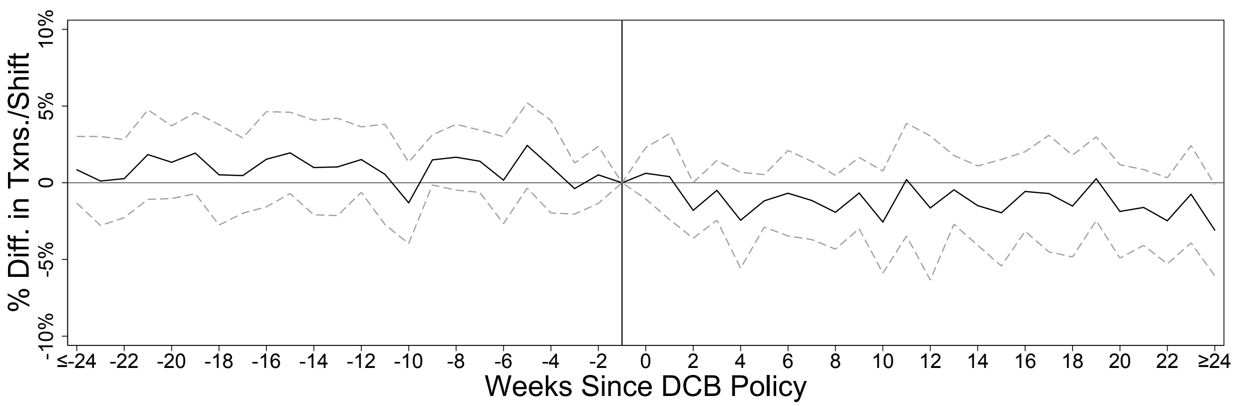
Note: Figure presents the full specification of event study Equation 3, with cashier and week-of-sample fixed effects and control variables for the average number of items scanned per transaction, average amount spent per transaction, the types of items purchased, and the weeks of experience of cashier c in store s, jurisdiction j, and week-of-sample w. The dependent variable is logged average transaction duration for cashier c in store s, jurisdiction j, and week-of-sample w, measured in minutes. Upper and lower 95% confidence intervals are depicted in gray, estimated using two-way cluster robust standard errors on policy jurisdiction and week-of-sample.

Figure 10: Effect of DCB Policies on Number of Transactions Completed per Shift (*Store-Week Averages*)

(a) Transactions per Shift



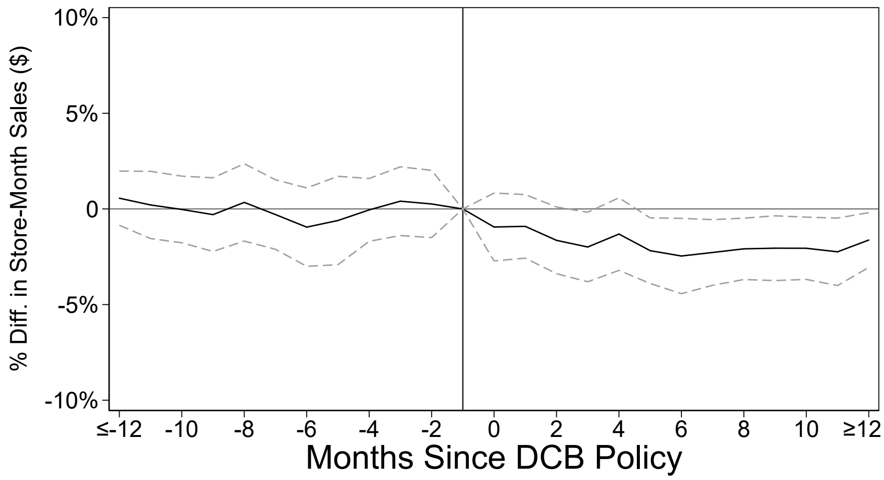
(b) Logged Transactions per Shift



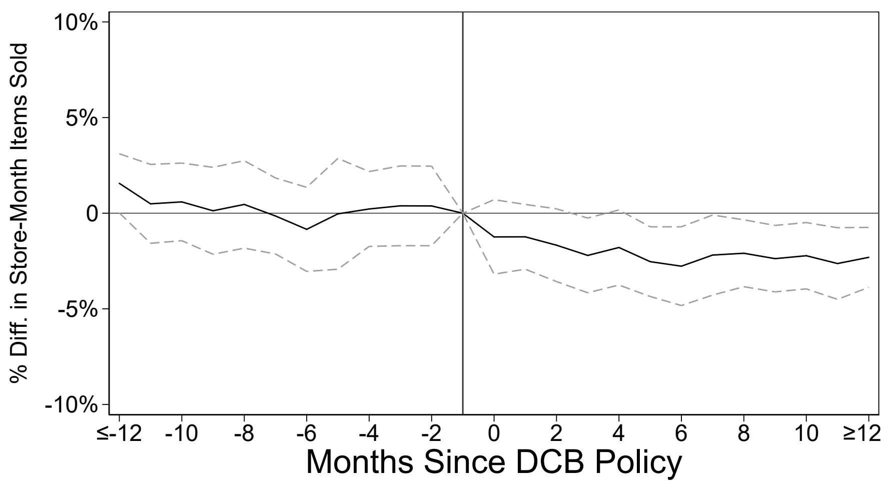
Note: The figure panels display the $\hat{\beta}_l$ coefficient estimates from event study Equation 1. The dependent variable is the number of transaction processed per 1:00-4:00pm shift in store s, jurisdiction j, and week-of-sample w, in levels (panel a) and in logs (panel b). Upper and lower 95% confidence intervals are depicted in gray, estimated using two-way cluster robust standard errors on policy jurisdiction and week-of-sample. The control variables include: average transaction size, average transaction expenditures, and the share of transactions purchasing each of the following items—produce, meat/seafood, dairy/refrigerated, frozen, bakery/deli, shelf-stable food, alcohol/tobacco, infant/toddler, floral department, and pet items—in store s, jurisdiction j, and week-of-sample w.

Figure 11: Effect of DCB Policies on Monthly Supermarket Sales
(*Nielsen Retail, Store-Month Averages*)

(a) Log Sales ($)



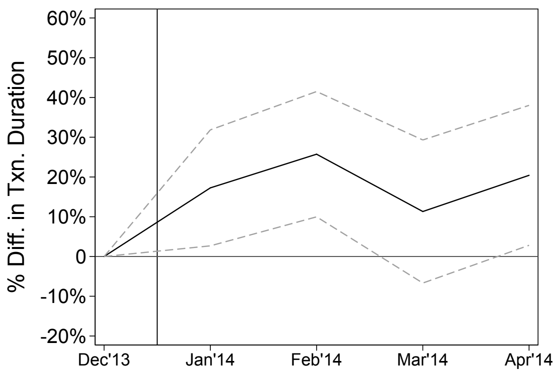
(b) Log Items Sold (#)



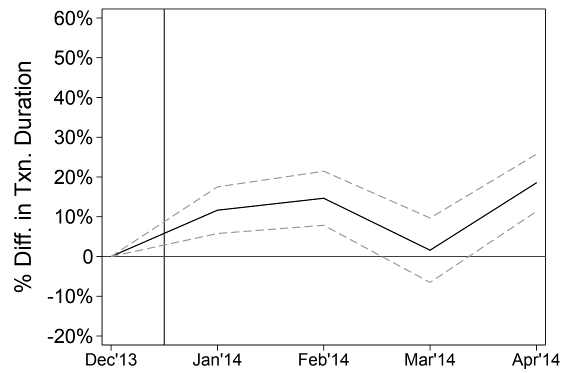
Note: The figures display the $\hat{\beta}_l$ coefficient estimates from the event study Equation 4 using the Nielsen Retail Data. The dependent variable in panel (a) is logged sales (measured in dollars), in store s, jurisdiction j, and month-of-sample m. The dependent variable in panel (b) is logged number of items sold in store s, jurisdiction j, and month-of-sample m. Upper and lower 95% confidence intervals are depicted in gray, estimated using cluster robust standard errors by jurisdiction and month-of-sample.

Figure 12: Effect of DCB Policies on Transaction Duration and Bagger Presence (*In-store vs. Scanner Data*)

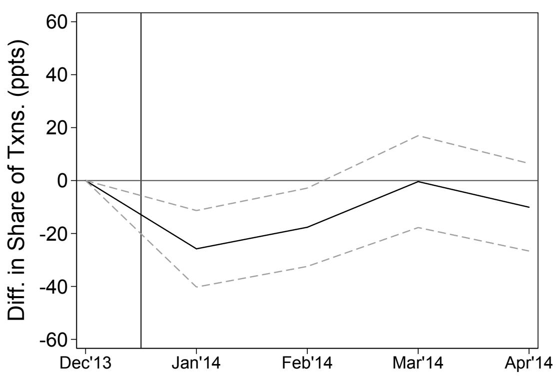
(a) Full-service Registers (*In-store Data*)



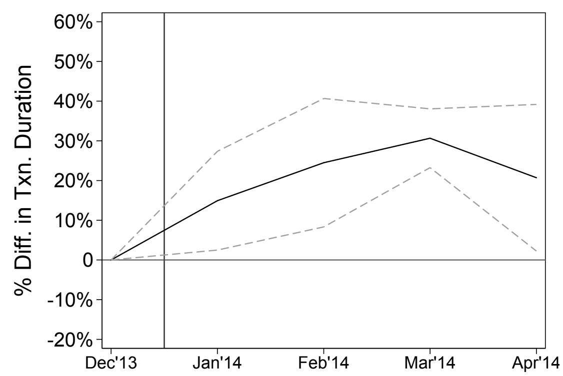
(b) Full-service Registers (*Scanner Data*)



(c) Bagger Presence (*In-store Data*)



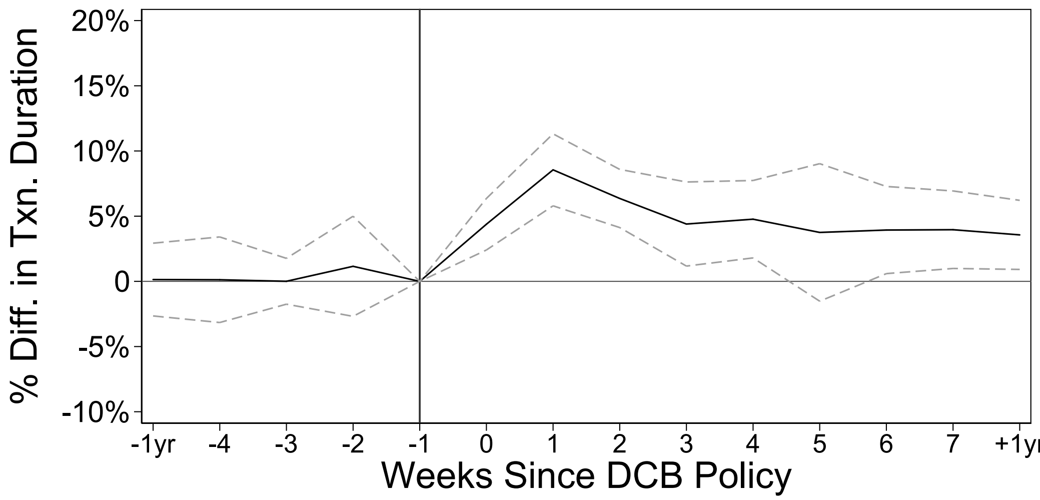
(d) Full-service Registers (*Discount Chain, In-store Data*)



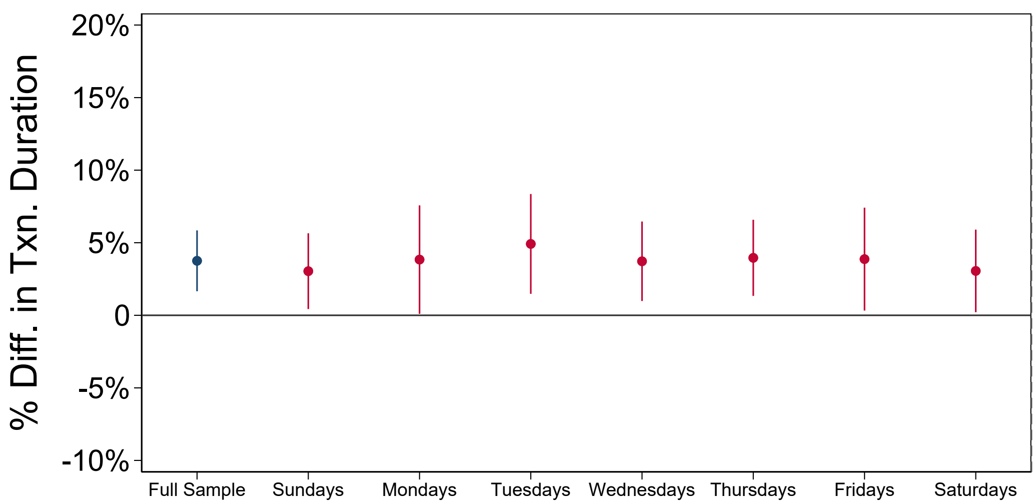
Note: The figure panels display the $\hat{\beta}_l$ coefficient estimates from the full specification of event study Equation 6. The dependent variable in panels (a), (b), and (d) is logged average transaction duration, measured in minutes, of transaction t in store s and day-of-sample d. The dependent variable in panel (c) is an indicator equal to 1 if transaction t had a bagger present. Panels (a), (c), and (d) use observational data collected in store while panel (b) uses scanner data. Panels (a), (b), and (c) use data from the main chain while panel (d) uses data from the discount chain. This analysis includes only transactions occurring at full-service registers, and not express or self-checkout registers. With the observational data in panel (a) and (d), the control variables include indicators for the gender and race of the person paying, whether there was a checkout interruption, and the register number. With the scanner data in panel (b), the control variables include the number of items scanned, the amount spent, the register number, and hour and cashier fixed effects. Upper and lower 90% confidence intervals are depicted in gray. Standard errors are calculated using error clustering at the store-day level.

Figure 13: Effect of DCB Policies on Transaction Duration using Washington DC Scanner Data (*Store-Week Averages*)

(a) Event Study, *full sample*



(b) Difference-in-differences, *full sample and by day-of-the-week*



Note: Panel (a) displays the $\hat{\beta}_l$ coefficient estimates from the full specification of event study Equation 7, whereas panel (b) presents the coefficients of the difference-in-difference form of Equation 7, estimated for the full sample and by day-of-the-week. The dependent variable is average logged transaction duration, measured in minutes, in store s and week-of-sample w. Upper and lower 95% confidence intervals are depicted in gray, estimated using two-way cluster robust standard errors on store and week-of-sample. D_{-5} equals one for all weeks w in December 2008 through February 2009 (i.e., a year before policy implementation). Similarly, D_8 equals one for all weeks in December 2010 through February 2011 (i.e., a year after policy implementation).