Giving Plastic Bags the Sack: The Hidden Costs of Changing Behavior

Rebecca Taylor*

JOB MARKET PAPER

January 15, 2017

Please click HERE for the most recent version

Abstract

Governments often enact policies to incentivize consumers away from behaviors with negative externalities, at the expense of consumer convenience. Understanding the non-monetary costs consumers face has implications for social welfare evaluation and policy design; however, quantifying these costs is not always feasible. In this paper, I precisely identify and measure a hidden time cost of an increasingly popular environmental policy aimed at altering behavior—the regulation of disposable carryout bags (DCB). Using variation in local government DCB policy adoption in California from 2011–2014 as a quasi-experiment, together with high-frequency scanner data from a national supermarket chain, I employ an event study design to quantify the effect of DCB policies on the wait and processing time of checkout services provided by supermarkets. I find that DCB policies cause a 3% increase in transaction duration, which persists over the entire sample period. Given the capacity constrained queuing system of supermarket checkout, the 3% slowdown of individual customers compounds into an even larger congestion externality—with DCB policies leading to an average additional 1.09 minutes of wait and processing time per customer. Aggregating to the state-level, a statewide DCB policy would cost Californians 25.8 million hours annually. This paper extends the literature on the hidden costs of changing behavior as the first i) to quantify the time cost of a policy change separately from other non-monetary costs, ii) to examine how this recurring cost evolves as behaviors and habits adjust to the policy, and iii) to focus on a policy and setting where capacity constraints determine whether retailers or customers bear the incidence of the time cost. The results have implications for the design of policy incentives, and show that ignoring time costs, as well as institutional constraints, may overstate the welfare gains from policy-induced behavioral change.

*Agricultural and Resource Economics; University of California, Berkeley. Phone: 513-600-5777 (cell). Email: becca.taylor@berkeley.edu. I am indebted to Sofia Berto Villas-Boas, Peter Berck, Hilary Hoynes, and David Zilberman for their continuous encouragement and guidance. I also thank Michael Anderson, Maximilian Auffhammer, Kendon Bell, Joshua Blonz, Fiona Burlig, Aluma Dembo, Meredith Fowlie, Daley Kutzman, Jeffrey Perloff, Louis Preonas, John Schindler, and Andrew Stevens as well as the participants of the UC Berkeley ERE seminar and the WEAI Graduate Student 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 project was conducted with generous support from the NSF Graduate Research Fellowship Program and the Giannini Foundation. All errors are my own.

1 Introduction

Governments often enact policies to incentivize consumers away from behaviors with negative externalities, at the expense of consumer convenience. For example, policies to combat airborne pollutants and congestion from driving—such as driving restrictions by license plate number (Davis, 2008), high-occupancy vehicle lanes (Kwon and Varaiya, 2008), and time-varying road pricing (Gibson and Carnovale, 2015)—incentivize consumers to spend time and effort in altering when and how they drive their vehicles. Energy efficiency subsidies (Allcott, 2016; Fowlie et al., 2015), garbage pricing (Fullerton and Kinnaman, 1996), and bottle return refunds (Beatty et al., 2007; Ashenmiller, 2011) encourage consumers to spend time and effort in conserving energy, reducing waste, and recycling. These policies illustrate that when the environment and consumer convenience are at odds with one another, policymakers "ask" consumers to trade convenience to benefit the environment. This begs the empirical question: what are the time, effort, and psychological costs consumers trade in changing their behavior? While the economic literature widely acknowledges the importance of non-monetary costs—for (i) improving policy design, (ii) conducting welfare analysis, and (iii) avoiding unintended consequences[1]—quantifying these costs is often infeasible, causing them to be easily overlooked. This is particularly true when the population is vast and heterogeneous and the behaviors to be altered are frequent but individually short-lived.

In this paper, I explore a hidden time cost of an environmental policy aimed at altering consumer behavior. Specifically, I examine how local government 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

[1] The importance of non-monetary costs in changing consumer behavior is acknowledged for three key reasons. First, policies often have greater success in changing behavior when the trade-offs consumers face are incorporated into the policy design. In the behavioral economics literature, numerous studies have shown how option defaults can be set recognizing the non-monetary 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). Second, 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. For example, with respect to the unintended consequence of driving restrictions by license plate number, Davis (2008) found that consumers circumvented the policy by increasing the total number of vehicles in circulation. Third, accurate welfare analysis requires a complete picture of a policy's costs, including the non-monetary costs paid by individual consumers (Allcott and Kessler, 2015).

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.[2] Using high-frequency scanner data from a national supermarket chain and variation in DCB policy adoption over time and space in an event study empirical strategy, this paper addresses two fundamental questions: 1) What are the time costs to individual consumers of policy-induced behavioral changes, and 2) How do these costs evolve as people learn and adapt their behavior?

Several features of supermarket checkout and DCB policies make them an interesting setting to study the time costs of changing behavior. First, food shopping is a common, frequent, and arguably necessary behavior. In the United States, consumers purchase the majority of food at grocery stores, supermarkets, and superstores (Taylor and Villas-Boas, 2016a), with the average American 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 8.7 billion hours in-store each year in the United States.[3]

Second, supermarket checkout is a setting where changes in the time spent in an activity can be directly, and precisely, identified and measured. With time-stamped, transaction-level scanner data obtained from a national supermarket, 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 Crispy Crunch Cereal at $3.79 each), and importantly, how much time the transaction took to complete. Unlike in previous 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 DCB clean-up, recycling, and landfilling costing local governments millions of dollars per year,[4]

[2] Retail stores pass the cost of disposable bags on to their customers in the overall price of groceries.

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

[4] Local governments are estimated to spend 1.1 cents per bag in collection, processing, and landfilling costs (Herrera Environmental Consultants, 2008). Given that approximately 100 billion plastic bags are consumed in

lawmakers across the country have adopted DCB policies to change how their constituents obtain food. From 2007 through 2016, approximately 242 local government DCB policies were adopted across 20 states and the District of Columbia.[5]

Fourth, while lawmakers acknowledge the trade-off between convenience and the environment in regulating bags,[6] these policies have been typically evaluated based on the magnitude of behavior change and litter reduction, and not full social welfare. Several studies have found DCB policies to be quite effective in altering consumer bag choices (Taylor and Villas-Boas, 2016b; Homonoff, 2016; Convery et al., 2007; Dikgang et al., 2012).[7] However, little is known about how these policy-induced behavioral changes affect the time and convenience of individual consumers. Given the extensive literature showing that shopping convenience impacts where and what people purchase to eat,[8] and the literature showing that consumers dislike and actively avoid long wait times (Katz et al., 1991; Tom and Lucey, 1995; Hirogaki, 2014),[9] it is important to understand the trade-offs between convenient behaviors and environmentally-friendly behaviors in food acquisition.

I hypothesize that DCB policies increase the duration of checkout through each of the three main inputs into the production function of supermarket checkout—namely, (i) bags, (ii) labor, and (iii) capital. First, DCB policies directly change the choice set of bags and their prices, and different bags vary in packing time. Second, to implement DCB policies, cashiers

---

the U.S. each year (Clapp and Swanston, 2009), municipalities nationwide spend $1.1 billion per year to manage plastic bags.

[5]Numerous major U.S. cities have adopted DCB policies, including San Francisco, Washington DC, San Jose, Seattle, Austin, Boulder, Los Angeles, Sante Fe, Chicago, Minneapolis, and New York City. For a list of DCB policies by city, county, and state, see: *Californians Against Waste.* Online, *accessed Sep. 5, 2016.*

[6]For instance, the City of Portland states, "Single-use plastic carryout bags may offer short-term convenience, but they have long-term costs. Not only do single-use bags require resources such as petroleum and natural gas to manufacture, their disposal presents a number of problems as well." (Online, *accessed Sep. 10, 2016*).

[7]Taylor and Villas-Boas (2016b) finds that a plastic bag ban coupled with a paper bag fee 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 the use of paper bags. Homonoff (2016) studies the impact of a plastic and paper bag tax that went into effect in Montgomery County, 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).

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

[9]Moreover, 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.*)

must 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).[10] Third, checkout lanes are optimized for single-use plastic bags, and the number of checkout lanes is optimized to handle checkout traffic during peak shopping hours. Importantly, checkout machinery is *fixed in the short-run*. During non-peak hours, if transactions are slower, retailers have the option to open more lanes to ease congestion at the cost of paying additional cashiers. However, during peak hours, retailers are constrained by their fixed checkout capital, and thus, slower transactions translate to increased checkout congestion and longer wait times for customers. 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 several 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 DCB policy adoption 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 DCB policies on checkout duration (my first research question) by comparing checkout duration at stores in jurisdictions with DCB 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 (my second research question). For the event study analysis, I design a subset of scanner data, selecting data from comparable treated and control stores across California between January 2011 and May 2014. In total, the dataset contains 9.3 million checkout transactions made during 1,047 peak shopping hours across 49

[10] Kahneman's (2011) proposes a the dichotomy between two modes of thought—"System 1" is fast, instinctive, and subconscious and "System 2" is slower, more deliberative, and more conscious. With economic incentive and regulations, policymakers are forcing people to switch from fast thinking habits (System 1) to slow thinking optimization (System 2). As people adapt to the policies, they may return to the speed and ease of System 1.

supermarkets.

My event study results reveal that DCB policies cause a 3% average increase in checkout transaction duration. I document heterogeneity in the policy effects by transaction size (i.e., the number of items purchased) and by whether a customer chooses to pay for paper bags at checkout, with the smallest transactions not paying the disposable bag fee experiencing no slowdown and the largest transactions paying for disposable bags experiencing a 10% slowdown. Surprisingly, even though I observe evidence of learning at the cashier level, this learning does not eliminate the slowdown from DCB policies, which persists over the entire sample period.

While 3% slower checkout durations (or roughly 3.6 seconds more per customer) may seem negligible, over the 11.9 billion grocery shopping trips made per year in the U.S., this time cost aggregates quickly. Moreover, shoppers experience both the slowdown of their own transaction and the slowdown of all transactions ahead of them in line. I find that DCB policies lead to a significant increase in checkout congestion during peak shopping hours, with 19 fewer customers processed per store per three-hour shift. Using a simple queuing theory model, 19 fewer transactions means each checkout queue is 1 customer longer on average. Using supplementary data, I provide suggestive evidence that the transactions lost during peak hours shift into the previously less busy shoulder hours, where stores have the capacity to open more lanes. Aggregating to the state level, the longer wait and processing time from DCB policies would cost Californians 25.8 million hours annually (≈$343 million). In comparison, the collection, processing, and landfilling of DCBs is estimated to cost Californian taxpayers $154 million per year.[11] This taxpayer estimate does not include the environmental cost of plastic marine debris. Therefore, while the aggregate time cost I estimate exceeds the amount currently paid by Californians in managing plastic bags, it might not exceed the long run environmental costs of plastic in oceans and waterways.

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

[11] Author's calculations, given that Californians are estimated to consume 14 billion plastic bags per year (CA Senate Rules Committee, 2014) and that DCB collection, processing, and landfilling is estimated to cost taxpayers 1.1 cents per bag (Herrera Environmental Consultants, 2008).

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 supplementary 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 in the DC metropolitan area, I again observe checkout slowdowns due to the policy change; however, unlike the California bag bans, the slowdown from the bag tax lessens significantly over time.

This paper contributes to an emerging literature on the hidden costs of changing consumer behavior.[12] To my knowledge, I am the first (i) to quantify the time cost of a policy change separately from other non-monetary costs, (ii) to examine how this recurring cost evolves as behaviors and habits adjust to the policy, and (iii) to focus on a policy and setting where capacity constraints determine whether retailers or customers bear the incidence of the time cost. My results also relate to the literature on congestion and waiting—which concludes that people place a higher value on time spent waiting than they do on the same amount of time in other circumstances (Maister, 1985; Larson, 1987; Small and Verhoef, 2007; Abrantes and Wardman, 2011)—and has implications for policies where governments intervene to protect citizens from their own choices. 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, especially in settings where consumer behaviors are connected through queuing systems. While often challenging to measure, and thus easy to overlook, quantifying these costs is vital for accurate welfare analysis and improved policy design.

The remainder of the paper is organized as follows. Section 2 describes the setting, empirical design, and data. Section 3 describes the event study regression model. Section 4 presents the main results. Section 5 rules out alternative mechanisms behind the transaction slowdown.

[12] Just and Hanks (2015) model the hidden emotional costs of command-and-control policies and argue that ignoring emotional responses to policy change may cause significant deadweight loss. Allcott and Kessler (2015) evaluate the welfare effects of social comparisons in reducing energy consumption and show that ignoring the time, comfort, and psychological costs of the intervention overstates the welfare gain of the program by a factor of five. In a similar vein, Damgaard and Gravert (2016) study the annoyance cost of a nudge intervention and show that when not accounting for the hidden costs of reminders, the average welfare effects are overstated by a factor of ten.

Section 6 uses three supplementary datasets to explore the external validity of the results and perform robustness checks. Section 7 discusses the broader impacts of the time costs of DCB policies. Section 8 concludes.

2 Setting, Research Design, and Data

2.1 Background on Disposable Carryout Bags and Regulations

When first invented, plastic carryout bags were considered quite the 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 and waterways, where, due to their durability and water-resistance, they last for a long time. While the majority of single-use plastic bags are landfilled or littered, even when properly recycled, they can clog the machinery used to sort materials.

Each year Americans consume approximately 100 billion single-use plastic 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, 2008), which aggregates to municipalities nationwide spending $1.1 billion per year. This clean-up cost estimate does not include the environmental costs of plastic marine debris. 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, which can be consumed by fish, turtles, and whales that mistake them for food.[13]

Given the environmental and clean-up costs of DCBs, lawmakers across the country are turning to policies to regulate DCBs. As of December 2016, approximately 242 local government

[13] A survey of experts, representing 19 fields of study, rank plastic bags and plastic utensils as the fourth severest threat to sea turtles, 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 (CA Senate Rules Committee, 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 July 26, 2016*).

DCB policies had been adopted across 20 states and the District of Columbia.[14] DCB policies prohibit retail stores from providing customers with "free" bags at checkout, with the goal of curbing the consumption of single-use bags and encouraging the use of reusable bags. These policies use one or both of the following policy tools to alter consumer behavior: (1) bag bans—command-and-control approaches to regulate behavior directly (i.e., quantity regulations), and (2) bag fees—market-based approaches to incentive consumers to change their own behavior (i.e., price regulations).

California provides a rare quasi-experiment for analyzing the effects of DCB policies on checkout duration and learning. In California, DCB policies ban retail food stores from providing customers with disposable 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, 82 DCB policies were implemented in California, covering 111 city and county jurisdictions and roughly one third of California's population.[17] This local legislative momentum culminated with the nation's first statewide plastic bag ban, which was voted into law on November 8, 2016.[18]

Figure 1 maps the implementation of DCB policies at four points in time. 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.[19] This figure highlights the fact that DCB policies

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

[15] While the vast 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 bans over fees? California 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 Appendix Table A.1 for a list of California DCB policies and implementation dates from 2007 to 2014.

[18] While a statewide ban on plastic bags passed the California state legislature and was signed into law by the governor on September 30, 2014, opponents secured enough signatures to put the ban to a public referendum. On November 8, 2016, Californians voted and passed the Plastic Bag Ban Referendum (Proposition 67) by a margin of 52.9% (yes) to 47.1% (no).

[19] Similar to other local government waste regulations, DCB policies may be implemented by city councils (for

have varied greatly across both implementation dates and locations. My event study empirical strategy exploits the variation in DCB policies across time and space from this quasi-experiment to explore how DCB policies influence checkout duration and learning.

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 national supermarket chain.[20] 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 and control, using the database of DCB policies I constructed for California.[17] As a first step in ensuring that control stores are good counterfactuals for treated stores, I dropped all stores in counties where no DCB policies had been adopted yet. 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, affecting 8 of the control stores. Furthermore, after receiving the data I decided to drop 11 stores which experienced either closure, remodeling, or policy differences, as these events could confound my checkout productivity measures.[21] Thus,

incorporated areas), county boards of supervisors (for unincorporated areas), and county waste management authorities (for entire counties with opt-out options for incorporated areas).

[20] 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.

[21] Of the 11 dropped stores, 3 stores closed before (or soon after) the end of the sample period, 5 stores were remodeled to add self-checkout registers, 2 stores were sold to a different company, 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 8 of the dropped stores were in the treated group and 3 were in the control group, I lose 21%

in the final sample I have 49 stores—30 treated and 19 control—across 42 policy jurisdictions (i.e., 42 incorporated city and unincorporated county 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.[22] 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, a little over half of the stores have pharmacies and coffee bars, and 10% or less have gas stations, juice bars, and sandwich counters. Roughly 50% 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 across treatment groups.[23] Once again, none of the variables are statistically different across treatment groups. 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, a

of the treated stores and 14% of the control stores.

[22] Jurisdictions decide when DCB policies will be operative, not the stores in a jurisdiction. The operative date is specified in a jurisdiction's ordinance document (i.e., bill) which is passed and adopted into law. Examining the ordinance documents of all 111 jurisdictions in California that implement DCB policies between 2007 and 2014, I find that 21% of jurisdiction specified January 1 as the operative 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. Implementation dates vary across all days of the week. Importantly, while operative dates were not randomly chosen, the dates were also not selected in a systematic way across all jurisdictions which would bias the results.

[23] As mentioned above, these variables were created by Gicheva et al. (2010) using 2000 US Census data for each store's census block-group.

greater share of White residents and a lower share of Other race and Multirace 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 regions.

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 weekend afternoon hours because the retailer cited them as peak shopping hours in their stores.[24] Having peak hours assures that transactions in the scanner data occur back-to-back, with little or no downtime in between. During these hours, the dataset includes every individual item purchased or returned at each store. In total, I have 1,047 hours of data across 49 stores, for approximately 127 million items scanned and 9.3 million transactions.[25]

2.3 Outcome Variables

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 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.[26]

My main measure of checkout productivity 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

[24] Before pulling the scanner data, I asked the retailer for their peak hours. I verified the hours they provided with Google store hour data. Importantly, while I find that 1:00-4:00pm on weekends are peak shopping hours, they are not the only peak hours in a week. Additional peak hours include 9:00am-5:00pm on weekends and 3:00-6:00pm on week days.

[25] 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 56 transactions with more than 250 items scanned, as these were outliers.

[26] 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.

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.[27,28] My second measure of checkout productivity is: *Transactions per Shift*—the number of transactions completed in a store per 1:00-4:00pm weekend shift.

Table 2 presents transaction-level summary statistics for 2011, which predate all DCB policies in my sample. Transactions are separated by the register type in which they occurred—1) full-service, 2) express, and 3) self-checkout.[29] Overall, Table 2 indicates that treated and control stores have balanced transaction-level characteristics in the pre-period. At treated stores, the average transaction at a full-service register takes 2.01 minutes to complete, comprises of 19.58 items, and costs $57.70.[30] The average transaction at an express register takes 1.49 minutes to complete, comprises of 8.65 items, and costs $26.15.[31] 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.

To better understand checkout productivity at the store level, Table 3 reports average store-shift characteristics in 2011 for treated and control stores. The table separates the summary statistics for stores with self-checkout (in the top panel) from stores without self-checkout (in the bottom panel), since checkout at these stores is inherently different—i.e., stores with self-checkout are systematically larger in square footage than those without.[32] In the pre-policy

[27] In Section 5, I verify this assumption using observational data collected in-store at checkout.

[28] 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 1.97 percent of transactions.

[29] 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.

[30] 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.

[31] Transactions in express lanes are statistically different in treated and control stores, with express transactions in treated stores being larger in size and longer in duration than in control stores.

[32] When the stores are ranked by size, all 13 stores greater than 53,000 ft^2 have self-checkout, while all 20 stores less than 41,000 ft^2 do not. To further understand how stores with and without self-checkout registers differ, I replicate Table 3 by self-checkout status. While stores without self-checkout are 10 years older and less likely to have many of the departments that stores with self-checkout have (such as coffee and juice bars), stores with

period, treated stores with self-checkout process approximately 705 transactions, $29,000 in sales, and 9,800 items per 1:00-4:00pm weekend shift. While full-service registers process only 50% of these transactions, they process over 70% of the items scanned and money spent. The smaller stores without self-checkout process 447 transactions, $15,400 in sales, and 5,200 items per shift. Yet once again, full-service registers at stores without self-checkout process 50% of the transactions and 70% of the items.

Table 3 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 how stores can react to increases in checkout duration and congestion due to a policy shock in the short run.

3 Empirical Model: Event Study Design

I estimate the causal effect of DCB policies on checkout duration using an event study design. This approach can be thought of as unpacking a difference-in-differences 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 the store-week level and employ the following event study regression model:

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

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

and without self-checkout are balanced across most demographic characteristics.

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.[33] 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 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

[33] 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 30 treated stores have at least fifteen weeks in the post-policy period, only 25 stores have thirty weeks, only 20 stores have sixty weeks, only 10 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.

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.

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 Results

4.1 Average Effects of DCB Policies on Transaction Duration

The figures in this section present the results from the estimation of event study Equation 1, where the $\hat{\beta}_l$ point estimates and 90% confidence intervals are displayed graphically.[34] Unless specified otherwise, I cluster the standard errors two ways—by jurisdiction (42) and by week-of-sample (177)—to allow for spatial and temporal correlation in the data.[35]

In Figure 3, the transaction-level scanner data are averaged to the store-week level, for a total of 8,673 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 treated and control stores l weeks from the DCB 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) alcohol and tobacco, (b) floral department items, (c) fresh meat and seafood, (d) fresh produce, (e) pet items, and (f) baby items. In Section 5, I verify

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

[35] 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.

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 $\hat{\beta}_0 = 0.033$, which means that the average transaction duration at treated stores is 3.3% 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 $\hat{\beta}_{24} = 0.023$. The $\hat{\beta}_{24}$ coefficient indicates that for all weeks in which it has been 24 or more weeks since DCB policy implementation, transactions at treated stores remain 2.3% longer than at control stores.[36] The majority of the post-policy $\hat{\beta}_l$ coefficients are significantly greater than zero at the 10% significance level. Importantly, none of the pre-policy $\hat{\beta}_l$ coefficients are 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,[37] however, I cannot reject that $\hat{\beta}_0 = \hat{\beta}_{24}$.[38] 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, I do not find evidence of transaction durations returning to pre-policy levels over time as customers, cashiers, and stores grow accustomed to the DCB policies.

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 ten.

In addition to the event study model in Equation 1, I estimate the following difference-in-

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

[37] $F(24, 41) = 3.63$, p-value = 0.000.

[38] $F(1, 41) = 0.98$, p-value = 0.329.

differences (DID) model:

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

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 (2) of Table 4. I estimate $\hat{\beta}_D = 0.030$ (p-value = 0.000), which is consistent with the event study results in Figure 3.

What are the implications of a 3% slowdown? A 3% increase in transaction duration means that the median 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,[39] meaning 3.6 seconds per grocery shopping trip would equal 11.9 million additional grocery shopping hours per year. That is equivalent to 1,358 years.

4.2 Heterogeneity by Transaction Size

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 characteristics of the transactions. First, I examine whether the effects of DCB policies on transaction duration are heterogeneous by transaction size—i.e., the number of items scanned in a transaction. To do this, I split the roughly 9.3 million transactions in my sample into four equal size quartiles: Q1=transactions with 3 or fewer scans, Q2=transactions with 4-8 scans, Q3=transactions with 9-18 scans, and Q4=transactions with 19 or more scans.

To understand how the size of transactions interacts with the effect of DCB policies, I estimate Equation 1 by size quartile (i.e., I average the transaction-level scanner data to the store-by-week-by-size-quartile level). Figure 5 presents the results of the full model specification estimated separately for each size quartile in ascending order. In this figure, the outcome

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

variable, Y_{sjw}, is again logged average transaction duration, which means the $\hat{\beta}_l$ point estimates measure the percent difference in transaction duration between treated and control stores l weeks from the DCB policy implementation. In panels (a) and (b), I find no statistically significant slowdown in transaction duration due to the DCB policies for the smallest two quartiles of transactions. Conversely, in panels (c) and (d), I find strong evidence that the policies cause slowdowns for transactions in the largest two quartiles. During the first week of the policies, Q3 transactions are 4.3% longer ($\hat{\beta}_0 = 0.043$) than prior to the policy. The $\hat{\beta}_l$ coefficients decline over time with $\hat{\beta}_{24} = 0.027$. For Q4 transactions, the slowdown starts with $\hat{\beta}_0 = 0.051$ and continues through the end of the sample with $\hat{\beta}_{24} = 0.032$. Using Wald tests for Q3 and Q4, I find mixed evidence that the slowdown for larger transactions lessens over time. While I cannot reject that $\hat{\beta}_0 = \hat{\beta}_{24}$ at the 10% significance level, I can reject that all coefficients in the post-policy period are the same as one another.[40]

Together these results suggest that the impact of DCB policies increases with the transaction size.[41] In other words, the policies do not impose a fixed time cost but instead the time cost depends on the number of items purchased. Why do DCB policies affect transactions of various sizes heterogeneously? One hypothesis is that transactions of different sizes choose different types and quantities of carryout bags. To understand how bag choice varies by transaction size, I estimate Equation 1 separately for each transaction size quartile, 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.[42] Figures 6 and 7 present the results.

As one would expect, in all panels of Figure 6 I find a sharp and permanent increase in the share of customers purchasing paper bags which is contemporaneous with DCB policy implementation. Since paper bags were available but not sold before the DCB policies, this figure

[40] If I estimate the DID model in Equation 2 instead of the event study model in Equation 1, I find slowdowns due to the DCB policies for the largest three quartiles ($\hat{\beta}_D^{Q2}$=0.019, $\hat{\beta}_D^{Q3}$=0.032, and $\hat{\beta}_D^{Q4}$=0.047), all statistically different from zero at the 10% level. Since larger transactions have longer checkout durations to begin with, this translates to Q2 transactions being 2 seconds slower, Q3 transactions being 4 seconds slower and Q4 transactions being 8 seconds slower on average.

[41] To formally test whether the event study results differ by transaction size, I perform a Chow Test (Chow, 1960)—comparing the residual sum of squares from the separate transaction quartile regressions to the residual sum of square of the whole sample. I calculate an F-statistic of 374.315 and can thus reject, at the 1% significance level, the null hypothesis that transaction size has no impact on the effects of DCB policies.

[42] It is important to note that in the scanner data I see whether or not a transaction pays the paper bag fee, but not how many paper bags are purchased.

reassures me that I have the correct timing of the DCB policy implementation. Additionally, I find that the share of transactions where consumers purchase paper bags increases with transaction size, with approximately 6% of Q1 transactions, 24% of Q2 transactions, and 32% of both Q3 and Q4 transactions choosing to purchase paper bags in the first week of the policy. In each panel, these shares decrease by roughly 5 percentage points over time.[43]

For reusable bags in Figure 7, 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 DCB policies.[44] As with paper bags, the share of transactions choosing to buy reusable bags increases with transaction size, with less than 1% of Q1 transactions, 2% of Q2 transactions, 7% of Q3 transactions, and 12% of Q4 transactions choosing to purchase reusable bags in the first week of the policy. However, these increases quickly retreat, and by eight weeks after the DCB 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 couple weeks of the policy.

4.3 Heterogeneity by Both Transaction Size and Paper Bag Choice

To explore how transaction size and customer bag choice interact to influence the effects of DCB policies on transaction duration, I estimate Equation 1 by size quartile and by whether a transaction purchased a paper bag (i.e. the transaction-level scanner data are averaged to the store-by-week-by-size-quartile level for those that purchase paper at treated stores in the post-policy period and for those that do not). Figure 8 presents the results, with transactions not purchasing paper bags on the left and transactions purchasing paper bags on the right.[45]

I find a stark difference in the effect of the policies between transactions with and without

[43] Since DCB policies stipulate that customers using food assistance program benefits—such as Supplemental Nutrition Assistance Program (SNAP) and Women, Infants, and Children (WIC) benefits—may obtain paper bags without paying a bag fee, the shares reported in Figure 6 are a lower bound for the share of customers obtaining paper bags at checkout. Out of 38.8 million Californians, 4.4 million received SNAP benefits in 2015—roughly 11 percent of Californians ("Just the Facts: The CalFresh Food Assistance Program." *Public Policy Institute of California*. Online, *accessed May 17, 2016*). Thus the scanner data may miss a sizable chunk of paper bag use. Taylor and Villas-Boas (2016b) use observational data collected in-store and find a higher share of transactions obtaining paper bags, between 30 and 40 percent, when a California DCB policy is in effect. This discrepancy may also be due to cashiers occasionally forgetting to charge the fee.

[44] 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.

[45] I do not present the results for Q1 because so few of these transactions purchase paper bags.

paper bag purchases, especially in the larger quartiles. For Q2 transactions that do not pay the paper bag fee, I find no effect of the policies. For Q2 transactions that do pay the paper bag fee, the post-policy $\hat{\beta}_l$ coefficients are positive on average, but the majority are not statistically different from zero. The same is true for Q3 transactions that do not purchase paper bags. However, for Q3 transactions that do purchase paper bags, the $\hat{\beta}_l$ coefficients in the post-policy period are greater in magnitude (e.g., $\hat{\beta}_0 = 0.070$ and $\hat{\beta}_{24} = 0.053$) and statistically different from zero. For the largest transactions in Q4, the effects of DCB policies on transaction duration are even stronger. DCB policies lead to a 2-4% slowdown in transaction duration for Q4 transactions not purchasing paper bags, and to a 7-10% slowdown for transactions that do purchase paper bags.

Overall, these results suggest that customers purchasing paper bags experience larger slowdowns than those not purchasing paper bags of the same transaction size. 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.[46] Yet since I run the regressions within transaction size quartile and control for the average amount spent and types of items purchased, conditional on observables these estimates are quite suggestive that paper bag choice is a mechanism behind the slowdown. Which raises the following question: Is the slowdown caused by the additional action of paying for paper bags, or are paper bags simply slower to pack than other types of bags? Given that cashiers enter in the bag fee code once per transaction, no matter the transaction size, one might expect the percent change in transaction duration from entering the fee to be larger for smaller transactions than for larger transactions. This is not what I find here, where the larger transactions experience the larger percent changes. These results provide evidence that bag type is an important mechanism behind the persistent slowdown, with paying for paper bags additively slower than getting "free" plastic bags.

[46] In Appendix A.2, I show that paper bag use is positively correlated with income, transaction size, and purchasing more expensive items.

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 larger transactions and 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 what customers purchase?

Are the slowdowns in transaction duration driven by changes in what customers buy when the DCB policies go into effect? In Figure 9, 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 9 has a different outcome variables, Y_{sjw}, at the store-week 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.[47]

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 2.9 cents lower than at control stores ($\hat{\beta}_1 = -0.029$). While I do not observe the size or volume of items purchased, 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,[48] it is unlikely to be the mechanism behind the persistent slowdown in checkout transactions.

[47] 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.

[48] The average item in 2011 costs \$2.97, so a 2.9 cent drop in price is less than a 1% change.

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. I find no significant changes due to the DCB policies in the share of transactions purchasing these items.[49]

5.2 Mechanism 2: Do DCB policies alter where customers checkout?

Along with choosing how many and what type of groceries to buy, customers also choose at which register to queue. In Figure 10, I first estimate Equation 1 with the outcome variable being either the share of transactions in cashier-operated registers or in self-checkout registers. Second, to see whether stores open more registers in response to the DCB policies, I also estimate Equation 1 with the outcome variable being the average number of registers open for each of the register types.

The results in Figure 10 are organized with register share as the outcome variable in the left-hand side panels and the number of registers open as the outcome variable in the right-hand side panels. Overall, I do not find large changes in either the share of transactions by registers type, or in the number of registers open, that are contemporaneous with policy implementation. In panels (a) and (c), I find that after the DCB policies are implemented, the share of transactions completed at cashier-operated registers declines slightly over time, and similarly, the share of transactions completed at self-checkout registers increases slightly. In particular, $\hat{\beta}_{24} = 0.014$ in panel (c), indicating a roughly 1 percentage point increase in the share of transactions at self-checkout 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

[49] The difference-in-differences estimates can be found in Appendix Table A.4. While I do not find changes in purchasing behavior when looking at these broad categories, I do find statistically significantly increases in one subcategory—garbage bags. This "plastic bag leakage" is yet another unintended consequence of DCB policies. In future work I will compare environmental consequences of the decline of thin plastic checkout bags with the increase in purchases of other types of bags.

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.

In panels (b) and (d), I analyze whether stores alter the number of registers open when the DCB policies go into effect. In neither panel do I find statistically significant changes in the number of cashier-operated and self-checkout registers open. Therefore it does not appear that stores are opening more lanes to mitigate the effects of DCB policies on checkout duration. However, since my sample of transactions come from peak grocery shopping hours, where stores are operating near full register capacity, stores may be unable to open more lanes due to capital and labor constraints. In Section 6.1, 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.[50] 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 Appendix A.1, I similarly explore the sensitivity of the results at the customer level, with customer fixed effects.

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

[50] Bertrand et al. (2004) discuss issues with estimating difference-in-differences 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.

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:

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

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 11 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. Additionally, I 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. 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 3,876 cashiers across the 49 store. On average, 37 cashiers work the 1:00-4:00pm weekend shift per store over the 3.5 years of the sample, with a minimum of 23 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.

Figure 11 presents a slightly different pattern than the store-week analysis in Figure 3. First, at the cashier-week level I find that the slowdown in transaction duration began a week or two before the policy.[51] 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 11, $\hat{\beta}_0 = 0.052$, while in Figure 3b, $\hat{\beta}_0 = 0.038$. Third, there is stronger evidence of learning at the cashier level than at the store level, with the post-policy

[51] In Figure 11, 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.

$\hat{\beta}_l$ coefficients diminishing in size over time ($\hat{\beta}_{24} = 0.024$). 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 5% significance level, however, I cannot reject that $\hat{\beta}_0 = \hat{\beta}_{24}$.

5.3.1 Cashier Learning after DCB Policies vs. Learning after Starting New Shift

To further explore the extent of cashier learning, I examine how cashier learning after DCB policies compares to cashier learning when a cashier first starts working at a store in the 1:00-4:00pm weekend shift. In particular, I estimate the following model:
For $Treat_{csj} = 0$:

$$Y_{csjw} = \sum_{e=1}^{177} \eta_e E_{e,csjw} + \beta_x X_{csjw} + \alpha_{csj} + \delta_w + \epsilon_{csjw} \tag{4}$$

For $Treat_{csj} = 1$:

$$Y_{csjw} = \sum_{e=1}^{177} \eta_e E_{e,csjw} + \sum_{l=-24}^{24} \beta_l D_{l,jw} + \beta_x X_{csjw} + \alpha_{csj} + \delta_w + \epsilon_{csjw} \tag{5}$$

where Y_{csjw} is the logged average transaction duration for cashier c in store s, jurisdiction j and in week w, α_{csj} is a vector of cashier fixed effects and δ_w is a vector of week-of-sample fixed effects. $E_{e,csjw}$ is a dummy variable equaling one if cashier c appears in the sample in week w for the eth time (i.e., $E_{1,csjw} = 1$ for all weeks in which cashiers appear in the sample for the first time). $D_{l,jw}$ again is a dummy variable equal to 1 if jurisdiction j in week w enacted a DCB policy l weeks ago, with $l = 0$ denoting the week of implementation. $Treat_{csj}$ is a dummy variable equal to 1 if cashier c is in one of the 30 stores treated. The first week the cashiers are in the sample ($e = 1$) and the first week of the DCB policies ($l = 0$) are the omitted dummies.

Plotting the first twenty-four $\hat{\eta}_e$ estimates alongside the post-policy $\hat{\beta}_l$ estimates allows me to compare the learning curve from starting to work in the 1:00-4:00pm weekend shift versus the learning curve from working with a DCB policy in place. For graphing purposes, I estimate the model separately for treated and control stores, however the results do not change when I pool the sample.

Figure 12 presents the results. On the left side of the graph, I plot the $\hat{\eta}_e$ estimates. The estimates for cashiers at treated stores are depicted with red circles and the estimates for cashiers at the control stores are depicted with blue hollow circles. I find that $\hat{\eta}_2 = -0.022$ at the treated stores, which means that between the first and second week cashiers work the 1:00-4:00pm weekend shift, they become 2.2% faster in completing a transaction. $\hat{\eta}_3 = -0.033$ means cashier checkout speed is 3.3% faster in the third week of work than the first week. The quickening of checkout duration continues at a diminishing rate. By week 24 of working the 1:00-4:00pm shift, the reduction in speed ceases and cashiers remain approximately 10% faster than their first week. This pattern holds for cashiers at the control stores as well.

I fit these coefficients into the conventional from of a learning curve (Alchian, 1963; Argote and Epple, 1990):

$$T_N = T_1 * N^b \tag{6}$$

where T_N is transaction duration for the Nth week of working the 1:00-4:00pm shift, T_1 is transaction duration in the first week, and $b = \frac{ln(LearnRate)}{ln(2)}$ is the slope of the learning curve. I estimate $T_N = 1.904 * N^{-0.033}$ which corresponds to a learning curve rate of 97.7%. This means that transactions in the second week take 97.7% the time of the first week, and transactions in the fourth week take 97.7% of the second week and so on.[52]

On the right side of Figure 12, I plot the post-policy $\hat{\beta}_l$ estimates (i.e., $\hat{\beta}_1$ to $\hat{\beta}_{24}$). I find evidence of changes in checkout speed from the first week of DCB policies (the omitted $\hat{\beta}_0$) to the subsequent policy weeks in that most of the $\hat{\beta}_l$ estimates are negative. Moreover, after policy week ten, where $\hat{\beta}_{10} = -0.25$, the majority of $\hat{\beta}_l$ coefficients are statistically less than zero at the 10% significance level. Note, I am not estimating the effect of the policy compared to the pre-period as I did in the event study estimations in Figure 11. Instead, I am estimating the change from the first week of the policy.[53] Fitting these coefficients into a learning curve (Equation 6) reveals $T_N = 1.879 * N^{-0.008}$, where T_N is now transaction duration in the Nth week of the policy. This corresponds to a learning curve rate of 99.4%.

Comparing Figure 11 and the halves of Figure 12 suggests that cashiers do learn and get

[52] In comparison, the 1-year death rate for hospitals performing heart transplants follows a 79% learning curve and the production rate of aircrafts follows a 80% learning curve (Heizer and Render, 2013).

[53] This is similar to re-centering the estimates in Figure 11 so that β_0 lies on the x-axis instead of β_{-2}.

faster after DCB policies; however, this learning curve is much shallower than the learning curve when starting a new shift. Moreover, 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. This exercise also provides a framework for evaluating the magnitude of checkout productivity slowdown from the DCB policies. A 3% slowdown in checkout speed would be similar to switching a cashier who has worked 20 weeks with a cashier that has only worked 10 weeks, or switching a cashier that has worked 3 weeks with one that is just starting.

5.3.2 Within Customer Effects

I also explore the sensitivity of the results at the customer level. The important take-away from the customer level analysis, presented in Appendix A.1, 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. I also find that the heterogeneity results in Figure 8—where larger transactions purchasing paper bags experience greater slowdowns—replicate at the customer level. This is reassuring because, at the store-week level, splitting the transactions by whether a paper bag was purchased in the post-period meant that the treated customers in the pre-period were not necessarily the same as the treated customers in the post-period. With the customer level data, I split the customers at treated stores in four groups by whether they ever buy paper bags and by transaction size. I find that none of the treated household groups differ from the control households in the pre-period. Yet after the policy, treated customers with larger transactions and those that choose paper bags have longer transaction durations than control customers.

6 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 Section 6.1, I compare the effects of DCB policies on transaction duration using scanner data versus using observational data collected in-store. In Section 6.2, 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 Section 6.3, I analyze whether slowdowns occur under a different type of policy (i.e., a bag tax), in a different state and time period.

6.1 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. I 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. 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.[54] 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 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 has yet to adopt a DCB policy. The two control cities were chosen to match Richmond with respect to average demographic characteristics.[55]

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

[55] I designed the in-store data collection 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

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

$$Y_{tsjdm} = \sum_{l=-1}^{3} \beta_l D_{l,jm} + \beta_x X_{tsjdm} + \theta_{sj} + \delta_{dm} + \epsilon_{tsjdm} \tag{7}$$

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 13 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.[56] 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

policies go into effect. Please see Appendix A.3 for a more detailed description of the variables in the in-store data.

[56] 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.

comparable in size to the coefficients using scanner data.[57] These results are also consistent with the main event study results in Section 4, using the scanner data from 49 stores, in that I find a significant and persistent slowdown in transaction duration.[58]

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, 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.

6.1.1 Matching In-Store to Scanner Data

Since I have scanner and in-store data for the same days, hours and stores, I next 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.[59] Thus far, 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, I can 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 can reject that paper and plastic bag transactions take the same amount of time per item at the 5% significance level, I cannot reject

[57] In Appendix A.4, I estimate a difference-in-differences model using both the scanner and in-store datasets and find similar results as these event studies.

[58] However, the $\hat{\beta}_l$ coefficients are much larger using the three store sample. In Figure 13b, $\hat{\beta}_0 = 0.173$, which is 4 times larger than what was estimated in Figure 3b, where $\hat{\beta}_0 = 0.038$. 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).

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

that reusable and plastic bag transactions take the same amount of time per item. This suggests paper bags are a slower technology, taking roughly a second more per item than plastic and reusable bags.

6.1.2 Do Transactions Shift into Different Hours of the Day?

Using the scanner data from the three store sample, I can also explore whether transactions shift into the less busy hours around the peak hours. For my main analysis (Section 4), I only observe transactions that occur during the hours of 1:00pm to 4:00pm, and thus, I cannot test whether transactions shift into less busy hours of the day. In the scanner data from the three store analysis, while I have a much smaller sample of stores and days, I observe all transactions made between 11:00am and 7:00pm.

With these data, I estimate the difference-in-difference model in Equation 1 for each hour between 11:00am and 7:00pm, with the outcome variable being (a) the number of transaction completed in each hour (or groups of hours) and (b) the average number of registers open in each hour (or groups of hours). Figure 14 plots the $\hat{\beta}_D$ coefficients. In panel (a) I find evidence that while the number of transactions decreases during the peak hours of 1:00pm to 4:00pm, the number of transactions remains the same or increases in the less busy hours surrounding the peak hour.[60] This provides suggestive evidence that some transactions lost during peak hours, due to the slowdowns from DCB policies, are made up in different hours of the day. However, when I estimate the model summed to the store-by-date level (denoted as *All Day* in Figure 14), I still find a decrease in the number of transactions processed, though it is no longer statistically significant.

In panel (b), I examine whether treated stores alter the number of registers open by hour of the day. Similar to Figure 10, during the peak 1:00-4:00pm hours, I do not find evidence that treated stores change the number of registers open. During the less-busy hours of 11:00am-1:00pm and 5:00-7:00pm, which have fewer registers open on average in the pre-policy period, the $\hat{\beta}_D$ estimates are slightly larger. However, none of the estimates are ever statistically different from zero, potentially due to the small sample size.

The panels in Figure 14 provide suggestive evidence that DCB policies cause some transac-

[60] For instance, $\hat{\beta}_D = 19.218$ for the 12:00pm hour, $\hat{\beta}_D = -25.734$ for the 2:00pm hour, $\hat{\beta}_D = 22.225$ for the 5:00pm hour.

tions to shift into less congested hours of the day. Moreover, stores may react to DCB policies by opening more registers in hours shouldering the peak hours, when they are not constrained by register capacity.

6.2 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 collecting observational data at the chain used throughout this paper, I collected data at a discount chain within the same three treated and control California cities as describe in Section 6.1. 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 Section 6.1 with the observational data from the discount chain.[62] Comparing the event study results in Figure 13a and Figure 15, shows that DCB policies lead to increases in transaction duration, even at stores in a different retail chain. In fact, the percent slowdown in transaction duration at the discount chain is even larger than at the national chain. While this comparison is suggestive and not causal, policymakers might be concerned if DCB policies affect low-income shoppers more so than wealthier shoppers, or if DCB policies affect regional stores more so than national stores.

6.3 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.? To answer this question, I use scanner data from the same

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

[62] Please see Appendix A.3 for a more detailed description of the variables in the in-store data at the discount chain and Appendix A.4 for the results of a difference-in-differences analysis using these discount chain data.

supermarket retailer as above, 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 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. 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 my 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 all transactions during the peak hours of 3:00-5:00pm on weekends and 5:00-7:00pm on weekdays during these months. I select six stores within DC that are open without interruption between December 2008 and January 2011. I select an additional 12 stores, within a 20 mile radius of DC, that best match the DC stores in terms of building age and size and census block-group characteristics. This gives me six treated and twelve control stores.[64]

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

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

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

[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] Appendix Table A.7 presents the summary statistics for the treated and control stores in the pre-policy period. I find that treated and control stores are balanced across most store and demographic characteristics. With respect to transaction level characteristics, in addition to presenting the averages from the entire sample of transactions, I split the transactions in half by transaction size and present the averages for the smallest (less than 8 items scanned) and largest (8 or more items scanned) transactions. I drop transactions that occur in self-checkout registers because only 4 of the 18 stores have self-checkout during my sample period. Overall, I find that transactions in the treated DC stores during the pre-policy period take slightly longer, but have roughly the same size and expenditures as transactions in the control stores. At treated stores, the average transaction takes approximately 2 minutes to complete, comprises of 12 items scanned, and costs $35.

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 16 plots the results. Given the estimates I find from the California DCB policies—where the effect of DCB policies on transaction duration depends on the size of the transaction and whether or not a customer chose to pay the bag fee—I estimate the model for the entire sample (panels a and b) and by transaction size and disposable bag use (panels c to f). In panel (a) the outcome variable is in logs and in all other panels it is in levels.

I find strong evidence that the DC bag tax leads to slower transactions. In panel (a), during the first week of the policy transactions take 7.7% longer to complete. In panel (b), the 7.7% slowdown translates to 0.159 minutes more per transaction. This slowdown lessens substantially over time, with $\hat{\beta}_8 = 0.087$ (a year after the policy) nearly half the magnitude of $\hat{\beta}_0 = 0.159$. Unlike the California DCB policies shown in Figure 3, I can reject that $\hat{\beta}_0 = \hat{\beta}_8$.[65] Thus, even though the effects of the DC bag tax on transaction duration do not fully dissipate after a year, I find stronger evidence of learning than under the California bag bans.

Next I examine heterogeneity by transaction size and bag choice. For the smallest transactions that do not pay the bag fee (panel c), I find a small but not statistically significant slowdown in transaction duration after policy implementation. The opposite is true for small transactions that do pay the fee (panel d), with $\hat{\beta}_0 = 0.112$ minutes and $\hat{\beta}_8 = 0.089$ minutes. For the largest transactions (panels e and f), I find statistically significant slowdowns for both the transactions that pay the fee and those that do not, and these slowdowns diminish over time. For the larger transactions that do not pay for disposable bags (panel e), the slowdown peaks in the second week of the policy ($\hat{\beta}_0 = 0.092$, $\hat{\beta}_1 = 0.0239$, $\hat{\beta}_8 = 0.133$). For the larger transactions paying the fee (panel f), the slowdown peaks in the first week of the policy ($\hat{\beta}_0 = 0.242$, $\hat{\beta}_1 = 0.214$, $\hat{\beta}_8 = 0.085$). This suggests that the aggregate slowdown (panels a and b) was at first due to the newness of paying the fee (i.e., people were not expecting to pay for bags during the first week of the policy), but the persistent effects of the bag tax on transaction duration

[65] $F(1, 574) = 3.61$, p-value $= 0.0578$

result from the alternatives to paying for disposable bags.

In Figure 17, I estimate Equation 8 with the share of transactions paying the bag fee as the outcome variable, separately for the largest and smallest transactions. Previous research has found that the vast majority of customers paying the DC bag fee chose plastic over paper bags (Homonoff, 2016). I find that 65% of the largest transactions pay for a disposable bag in the first week. This drops to 56% by week 2 and 50% by week 3. A year after the policy, 45% of large transactions pay the fee. For the smallest transactions, 44% pay the fee in week 1 and 32% a year later. Thus it does appear that some shoppers were surprised by the policy in its first week and altered their bag choice behavior in subsequent weeks.

Comparing the event study results for the DC and California policies, I find evidence that the mechanisms behind the slowdown are not the same across policy tools, and that the slowdown is less persistent over time under a bag fee than a bag ban. The results in Figure 16 are similar to what I find in Figure 8 in one way—I do not find slowdowns due to either policy for the smallest transactions that opt not to pay for DCBs. However, the DC results differ from California in that the largest percent slowdowns a year after the policy do not occur for the larger transactions paying the bag fee. Instead, *both* large and small transactions paying the fee are 0.09 minutes slower than transactions at control stores a year after the policy. Therefore, when plastic bags have a fee, there is a fixed time cost of paying the bag fee which is independent of transaction size. Conversely, when plastic bags are banned and paper bags have the fee, there is also an additive time cost which scales with items purchased.[66] Adding this to the evidence that paper bags are a slower technology than plastic bags (Section 6.1), suggests that bag taxes may have lower time costs than plastic bag bans coupled with paper bag fees.

7 Discussion of Broader Impacts

7.1 Checkout Congestion, Queue Length, and Customer Wait Time

The results above indicate that DCB policies in California cause persistent 3% increases in the amount of time to checkout at supermarkets. How does this increase in processing time per

[66] Additionally, the DC results differ from California in that the larger transactions not paying the fee (panel e) experience slowdowns of similar magnitude as the larger transactions paying the fee (panel f). Since the DC policy predates the California policies in my sample by at least two years, this may be due to DC cashiers and customers having less experience with reusable bags when the policy went into effect.

transaction impact the amount of time customers spend in line waiting to checkout? During peak hours, when 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, I can use the scanner data and a simple queuing theory model to approximate how many more people are waiting in line due to DCB policies. First, I quantify the change in the number of transactions completed per store and shift by estimating Equation 2 with the outcome variable being the average number of transactions completed per three-hour weekend shift. If we assume that the arrival process to checkout does not change as a result of the DCB policies (an assumption I can relax later), then a decrease in the number of transactions processed would mean these "missing" transactions are still waiting in line to be processed. I then calculate the average increase in the number of transactions waiting to be processed as $\frac{|\hat{\beta}_D|}{2}$.[67]

The top panel of Table 4 present the $\hat{\beta}_D$ coefficients using the scanner data from the 49 stores across California, estimated in both levels and logs. In column (1), I find that stores process 19.261 fewer transactions per three-hour shift when DCB policies go into effect, which equals a 3.2% decrease in the number of transactions completed per shift.[68,69] In column (2) I use transaction duration again as the outcome variable instead of transactions per shift. Corroborating the event study results in Figure 3, I find that DCB policies lead to a 3.0% increase in transaction duration. Comparing columns (1) and (2), the 3% increase in transaction durations translates to a 3% decrease in the number of transactions processed, which means stores are not absorbing any of the slowdown from DCB policies during these peak-hour shifts.

I report the $\frac{|\hat{\beta}_D|}{2}$ estimate in the first row of the bottom half of Table 4 and then use it to calculate the additional number of customers in line *per register* either (i) given the average

[67]I divide $|\hat{\beta}_D|$ by 2 to get the average change in customers standing in line per store at any given moment during a 1:00-4:00pm shift, conservatively assuming the increase in line length is zero at the beginning of the shift and grows linearly to $|\hat{\beta}_D|$ by the end of the shift. This is conservative because the peak shopping hours extend before and after 1:00-4:00pm on weekends, making it likely that an increase in line length would have started before 1:00pm.

[68]In Appendix Table A.8, I explore whether the result in Table 4 column (1) varies by store characteristics. I find evidence that the decrease in transactions completed due to the policy change is greater for stores in census blocks-groups with a higher median income, a lower share of Asian residents, and a higher share of Black residents. This finding is consistent with the results in Appendix Table A.3—which showed paper bag use is positively correlated with income and negatively correlated with the Asian population share—and the result in Figure 8—which showed greater slowdowns for transactions choosing paper bags.

[69]In Appendix Figure A.2, I estimate this as an event study using Equation 1 and find similar results.

number of registers open in the post-policy period, or (ii) if all existing registers were open. In column (1), with an average 9.276 registers open in the post-policy period, the DCB policies cause a 1.038 transaction increase in queue length *per register open*. If instead all existing registers were open—the average of which is 10.735 registers—DCB policies would add an additional 0.897 transactions per queue. Therefore, the slowdown in transaction duration from DCB policies causes each checkout queue to be approximately 1 customer longer on average.

These are upper bounds for changes in queue length. At the other extreme, queue length may not change if customers not served during the 1:00-4:00pm shift decide to shop at a different (and potentially less crowded) time of the day/week, or, at a different store altogether. Not only is grocery shopping less convenient under DCB policies, restaurants and food-away-from-home establishments are 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 in the form of a convenience cost or the actual cost of bags), some of the 19 customers not served during the shift may have chosen to purchase food elsewhere. In future work I will empirically test whether DCB policies shift customers away from grocery shopping towards eating out.

7.1.1 Interpreting the Time Cost of Checkout Congestion

How would customers fare if each checkout line is 1 customer longer during peak hours? The median transaction in my sample in 2011 was approximately 2 minutes. 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.[70] Thus, if DCB policies cause the average queue to increase by 1 transaction during peak hours and half of all transactions occur during peak hours, this would translate to an average additional 1.09 minutes of wait and processing time ($1.09 = 2 * 1.03 * 0.5 + 2 * 0.03$). For busy customers, this is not a negligible wait time. Given the average grocery shopping trip on weekends lasts 44 minutes (Hamrick et al., 2011), a 1.09 minute longer wait translates to shoppers spending 2.5% more time in store per trip. An industry survey found that the average wait time in grocery shopping lines in 25 major cities

[70] "Grocery Shopping: Who, Where, and When." *Time Use Institute.* Oct. 2008. Online, *accessed Sep. 9, 2016.*

was 4 minutes, meaning DCB policies lead to a 27% increase in grocery shopping checkout wait time.[71]

Using half the average California hourly wage as the value of time—since grocery shopping often occurs during non-work hours when the opportunity cost of time is low[72]—1.09 minutes is worth $0.24. Aggregating to the state level, the total time cost of a statewide DCB policy for the 28 million Californian adults would be as much as 25.8 million hours annually (≈$343 million).[73] To get a sense of the magnitude of this time cost, I turn to the literature on the time savings of policies affecting traffic congestion. Anderson (2014) finds that Los Angeles public transit saves 114 million hours of traffic congestion delay each year. Foreman (2013) finds that a bridge toll price change in the San Francisco Bay Area saved 210,000 hours annually. Thus the time savings of issuing "free" bags at checkout in California is in line with the time savings from other policies that affect congestion.

To compare the time cost of DCB policies to the benefits of reducing plastic bag consumption, I use an estimate of how much taxpayers spend in collection, processing, and landfilling disposable bag waste, which is 1.1 cents per bag (Herrera Environmental Consultants, 2008).[74] Based on the in-store observational data of bag use at checkout, a statewide DCB policy would lead to the use of 4.7 billion fewer disposable bags per year. This would save $51.7 million in tax dollars annually.[75] However, this tax estimate does not include the environmental cost

[71] "Justice—Wait for It—on the Checkout Line." *Wall Street Journal.* Aug. 19, 2009. Online, *accessed May 30, 2016.*

[72] 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 California hourly wage is $13.29 ("May 2015 State Occupational Employment and Wage Estimates, California." *Bureau of Labor Statistics, U.S. Dept. of Labor* Online, *accessed May 28, 2016*).

[73] This monetary estimate is a lower bound for the cost of time as there is an extensive literature concluding that people place a higher value on time spent waiting than they do on the same amount of time in other circumstances (Maister, 1985; Larson, 1987; Small and Verhoef, 2007; Abrantes and Wardman, 2011). In quantifying the effect of public transit on traffic congestion, Anderson (2014) uses a delay multiplier of 1.8. Using this multiplier would bring the time cost estimates up to $617 million.

[74] The California Senate Rules Committee (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).

[75] In the in-store data, I find that the average customer uses 3.33 fewer disposable bags per transaction post-DCB policy. Given Californian adults make 1.42 billion grocery transactions per year, this equal 4.7 billion fewer disposable bags per year.

of plastic marine debris.[76] Thus while the aggregate time cost I estimate surpasses the taxes currently paid by Californians in cleaning up plastic bags, it might not surpass the long run environmental costs of plastic in oceans and waterways.

7.2 The Incidence of DCB Policies on Supermarkets

7.2.1 What Share of the DCB Policy Slowdowns Do Supermarkets Bear?

Stores could be hurt if productivity slowdowns lead to lost transactions and lower revenues. Given that the average transaction in my sample costs roughly $40 to the consumer, if all 19 customers not served per shift because of the DCB policies decide to purchase food elsewhere, stores would lose $760 in revenue each shift. This would be an extreme case. As suggested by Figure 14a, it is likely that some of these transactions spill into the next shift or are made up in other hours of the day/week. I do not have transactions for an entire day, and consequently, cannot measure whether total daily store revenue decreased due to the DCB policies. However, knowing that long waiting time is one of the factors that brings the most dissatisfaction to customers (Nomi, 2014; Katz et al., 1991; Tom and Lucey, 1995; Hirogaki, 2014), grocery stores could choose to open up more registers. To get back to the same level of transactions per shift from before the policies, store would need to open 0.323 more registers.[77] As shown in Figure 10, I do not find an increase in the number of registers during the peak 1:00-4:00pm hours. While stores cannot open more registers than they have available, stores can open more registers during the less busy hours shouldering the peak hours. If all 10,935 supermarkets and grocery stores in California opened 0.323 more registers during the 14 shoulder hours per week, this would cost $31 million in additional wages per year (at a $12 per hour wage rate).[78] In Figure 14b, I find weak evidence that stores do open more registers during the shoulder hours. In summary, while stores have the option of bearing the burden of DCB policies during non-peak hours, customers bear the burden of the policies during peak hours when checkout

[76] Jambeck et al. (2015) estimate that 1.7-4.6% of the plastic waste generated across 192 coastal countries around the globe is mismanaged and enters the ocean. Once in waterway, plastic bags do not biodegrade, but instead break into smaller pieces, which can be consumed by fish, turtles, and whales that mistake them for food.

[77] The number of additional registers needed to get back to the same level of transactions is calculated using estimates from Table 4 as follows: $0.323 = (\frac{573.142}{573.142-19.261} - 1) * 9.276$.

[78] This back-of-the-envelope calculation does not include the potential costs of having to retrain cashiers and baggers to pack varying types of bags.

capacity is constrained.

7.2.2 Supermarket Savings from No Longer Providing "Free" DCBs to Customers

It is important to consider the benefits of DCB policies to retail stores. In particular, DCB policies reduce the need to purchase disposable carryout bags to provide to customers. Standard single-use plastic bags cost retailers on average 3 cents each and paper bags cost 7 to 10 cents each.[79] In interviews, store managers list disposable bags as their fourth largest operating cost, after electricity, payroll, and credit card fees. Retail stores usually pass the cost of disposable bags on to their customers by incorporating them into the overall price of groceries. If stores do not adjust their prices down to account for the bags they no longer buy, stores could experience significant savings. Furthermore, the California DCB policies stipulate that grocery stores must sell paper bags for 10 cents, even if a store purchases bags at a lower price. This 10-cent fee is kept entirely by the store and is not a tax collected by the government. To get a sense of the magnitude of revenue stores would make on paper bags sales under a statewide policy, I use my in-store observational data on bag use at checkout. Given the average customer in the post-policy period buys 0.46 paper bags and that Californian adults make 1.42 billion grocery trips per year, supermarkets would make $65.3 million in revenue from paper bags sales annually.

7.2.3 Reoptimizing Checkout Lanes for Reusable Bags

Finally, supermarkets are optimized for single-use bags, with checkout registers and bagging areas designed for quickly dispensing these bags. It will take time to reoptimize checkout machinery for reusable bags—especially since retailers may want to see how many cities, counties, and states will pass DCB policies before investing in costly store remodels. One potential solution to reduce congestion from DCB policies would be the creation of "Reusable Bag Only" lanes (similar to express checkout lanes and HOV traffic lanes). Since the slowdown in transaction time is driven by those who choose paper bags, separating the paper bag users from the reusable bag users reduces the time externality paper bag users impose on others in line. As a thought experiment, I solve for the number of paper-bag-lanes and non-paper-bag-lanes to

[79] Bag cost estimates come from interviews with the store managers in my sample, but media articles also confirm these estimates ("Plastic Ban Means Higher Costs are in the Bag." *Crain's Chicago Business.* May 3, 2014. Online, *accessed Sep. 10, 2016*). The cost of paper bags depends on bag size and the presence of handles.

make the average wait time per lane proportional to the average transaction duration for the bag choice in that lane, using summary statistics from the post-policy scanner data.[80] This exercise reveals that a store could convert 22% of its full-service lanes to paper-bag-lanes and 78% to non-paper-bag-lanes and still process the same aggregate number of transactions per shift as if it had not sorted its customers.[81] However, while sorting customers may reduce the time externality of paper bags in theory, in practice sorting customers by bag choice might have consequences for the share of customers purchasing paper bags (e.g., separate bag lanes may increase stigma from purchasing paper bags) and for the duration of paper bag transactions (e.g., separate lanes may increase productivity though specialization), which would alter the results of this thought experiment.

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 3% increase in supermarket checkout duration. While I observe evidence of learning at the cashier level, this learning does not offset the slowdown from DCB policies, which persists over the entire sample period. In addition, I find the policy effects are heterogeneous by transaction size and by whether paper bags are purchased, with the largest transactions paying the paper bag fee experiencing a 10% increase in transaction duration.

The slowdown from DCB policies generates a congestion externality, where shoppers not only experience the slowdown of their own transaction, they also experience the slowdown of all transactions ahead of them in line. Using a simple queuing theory model, I calculate that DCB policies cause each checkout line to be roughly 1 customer longer on average during peak hours. This translates to an additional 1.09 minutes of wait and processing time per transaction on average. Aggregating to the state level, a statewide policy would cost Californian shoppers 25.8 million hours annually (≈\$343 million).

[80] The optimal ratio of paper-bag-lanes, L_p, and non-paper-bag-lanes, L_r, is $\frac{L_p}{L_r} = \frac{Share_p * TxnDur_p}{Share_r * TxnDur_r}$, where $Share_p$ and $Share_p$ are the shares of transactions choosing paper and non-paper and $TxnDur_p$ and $TxnDur_r$ are the average transaction durations by bag choice.

[81] The average wait time in non-paper-bag-lanes would be 2.5% shorter than in an unsorted lane because these transactions are 2.5% shorter than the unsorted average. The average wait time for paper-bag-lanes would be 10.2% longer than in an unsorted lane since these transactions are 10.2% longer than the unsorted average.

DCB policies are not the only recent change to cause longer lines at grocery stores. This research has implications for the roll-out of credit cards with chip technology to reduce credit card fraud.[82] An industry study found that using a chip card added 8 to 12 seconds per checkout transaction.[83] Similar to DCB policies, these slowdowns come from the chip readers being slower than swipe readers and from customers and cashiers needing to learn how the new technology works. The results of this paper imply that the benefits of increased security from chip technology should be compared against the time costs to consumers and retailers.

The policy implications of this paper are threefold. First, policies which incentivize customers to change their habits may have large non-monetary costs, and ignoring these costs overstates the welfare gains of such policies. Even though DCBs comprise a very small fraction of the monetary expenditures for food production, their role in reducing the time and effort necessary for healthy food production is not trivial. Second, not fully considering the institutional conditions and constraints of a policy setting can result in competing externalities. I show that when consumer behavior is connected through queuing systems, individually slower actions propagate into an even larger congestion externality. Third, the policy tool (i.e., bans vs. fees) matters with respect to the time costs. I find that policies which tax both plastic and paper bags have less persistent time costs than policies which ban plastic and tax paper, due to paper bags being the slower technology.

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. In future work I will examine whether and how consumers circumvent DCB policies in unintended ways—in particular, (i) by switching away from grocery shopping towards eating out more frequently and (ii) by increasing purchases of other types of disposable plastic bags.

[82] 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*.)

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

References

Abrantes, P. A. and M. R. Wardman (2011). Meta-analysis of UK Values of Travel Time: An Update. *Transportation Research Part A: Policy and Practice 45*(1), 1–17.

Alchian, A. (1963). Reliability of Progress Curves in Airframe Production. *Econometrica 31*(4), 679–693.

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

Allcott, H. and J. B. Kessler (2015). The Welfare Effects of Nudges: A Case Study of Energy Use Social Comparisions. No. w21671. National Bureau of Economic Research.

Anderson, M. (2014). Subways, Strikes, and Slowdowns: The Impacts of Public Transit on Traffic Congestion. *American Economic Review 104*(4), 2763–2796.

Argote, L. and D. Epple (1990). Learning Curves in Manufacturing. *Science 247*(4945), 920–924.

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.

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. http://www.ncpa.org/pdfs/st353.pdf, accessed Sep. 8, 2016.

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

Chow, G. C. (1960). Tests of Equality Between Sets of Coefficients in Two Linear Regressions. *Econometrica 28*(3), 591–605.

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.

Committee, C. S. R. (2014). Solid Waste: Single-use Carryout Bags. Senate Bill 270, Bill Analysis.

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.

Damgaard, M. T. and C. Gravert (2016). The Hidden Costs of Nudging: Experimental Evidence from Reminders in Fundraising. Working Paper.

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.

Foreman, K. (2013). Crossing the Bridge: The Effects of Time-Varying Tolls on Curbing Congestion. Dissertation, Department of Agricultural and Resource Economics, University of California, Berkeley.

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.

Gibson, M. and M. Carnovale (2015). The Effects of Road Pricing on Driver Behavior and Air Pollution. *Journal of Urban Economics 89*, 62–73.

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. Online, 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, U.S. Department of Agriculture, Economic Research Service.

Heizer, J. and B. Render (2013). *Operations Management, 11th Edition*. Pearson Prentice Hall.

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

Hirogaki, M. (2014). Service Quality Management in Retailing: Theory and Empirical Analysis of Customer Satisfaction with Checkout Services. *International Journal of Trade, Economics and Finance 5*(3), 270–276.

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.

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.

Kwon, J. and P. Varaiya (2008). Effectiveness of Californias High Occupancy Vehicle (HOV) System. *Transportation Research Part C 16*(1), 98–115.

Larson, R. (1987). OR Forum—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.

Nomi (2014). Why Customers Leave and How to Keep Them. White paper, Nomi/Brickstream Corporation. Online, accessed Mar. 3, 2016.

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.

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. *Marin Policy 65*, 107–114.

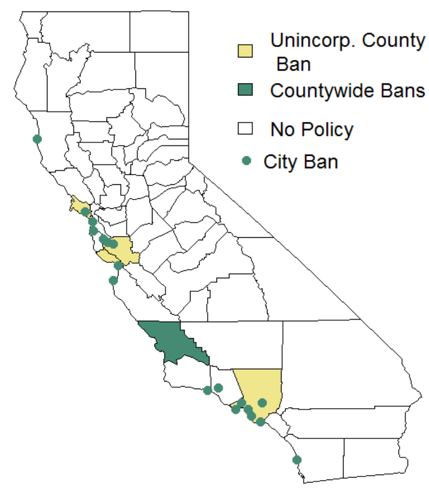
Yaktine, A. and J. Caswell (2013). *Supplemental Nutrition Assistance Program: Examining the Evidence to Define Benefit Adequacy*. National Academies Press, Washington, DC.

9 Figures

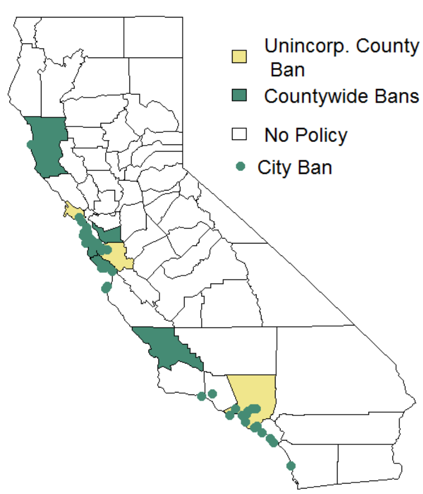
Figure 1: California Disposable Carryout Bag (DCB) Policies over Time

(a) Policies Implemented Before 2012

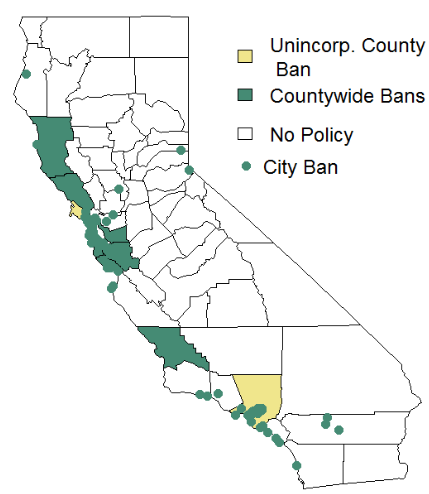
(b) Policies Implemented Before 2013



(c) Policies Implemented Before 2014

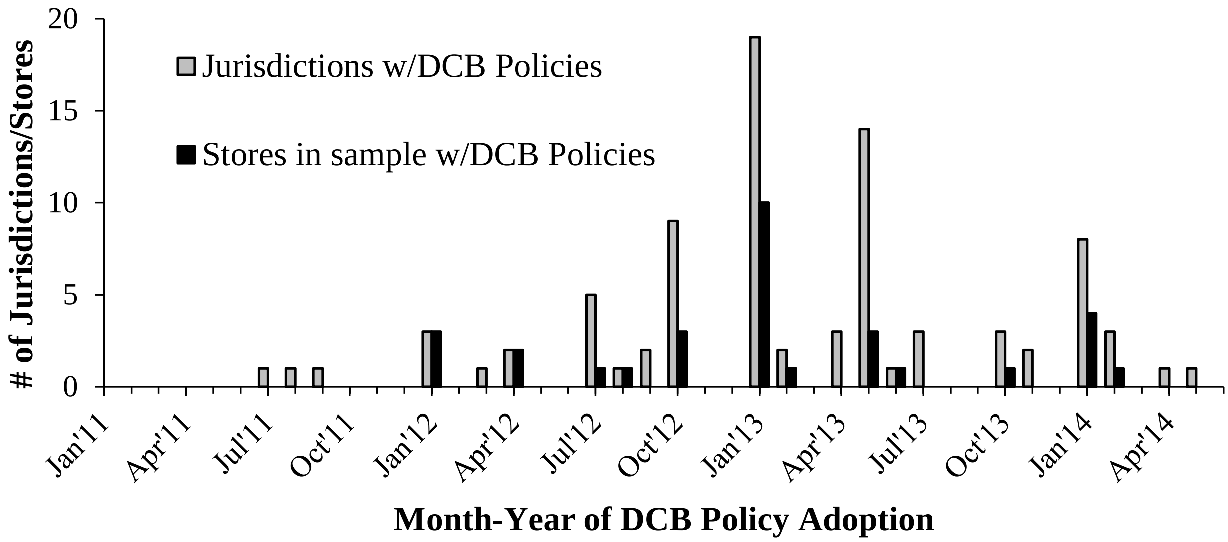


(d) Policies Implemented Before 2015



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.

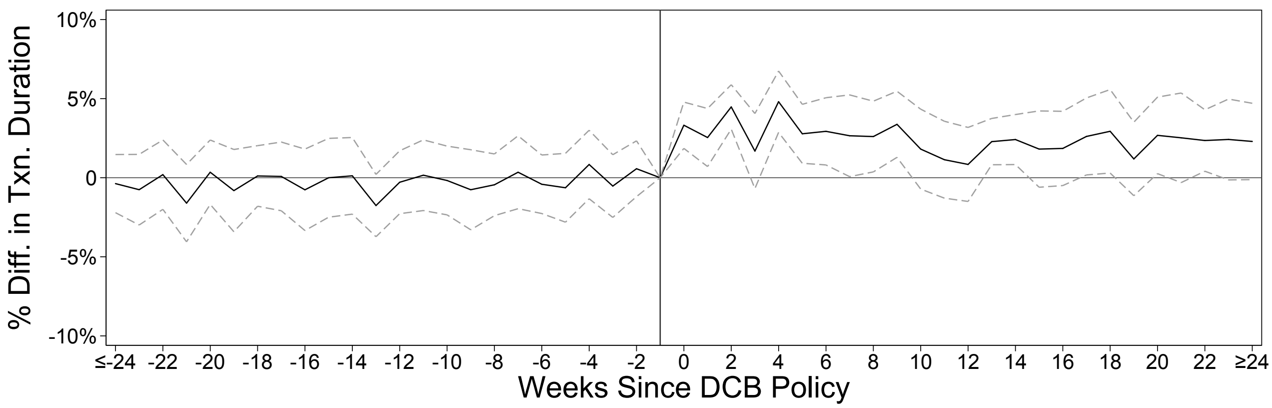
Figure 2: Number of California Jurisdictions Implementing DCB Policies by Month and Year



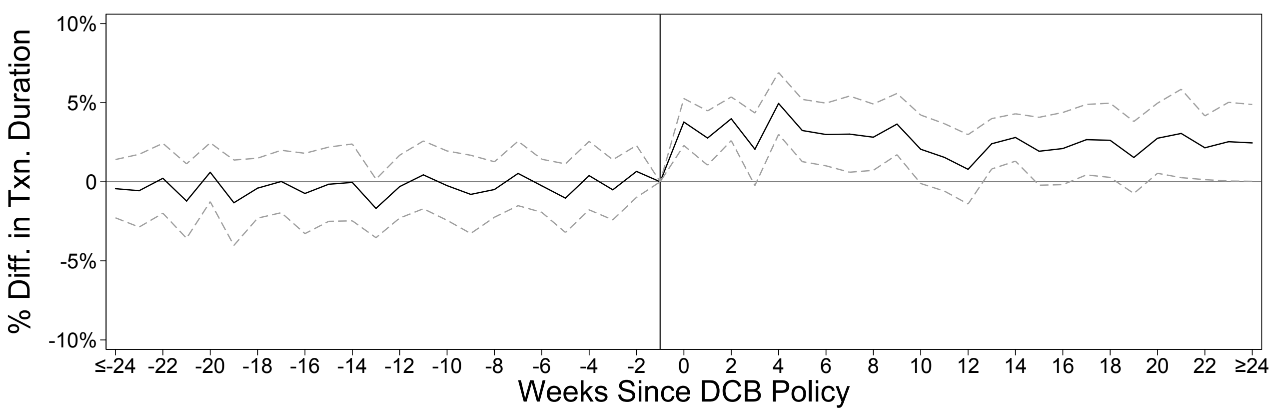
Source: Author's calculations.

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

(a) Logged Transaction Duration—Without Control Variables

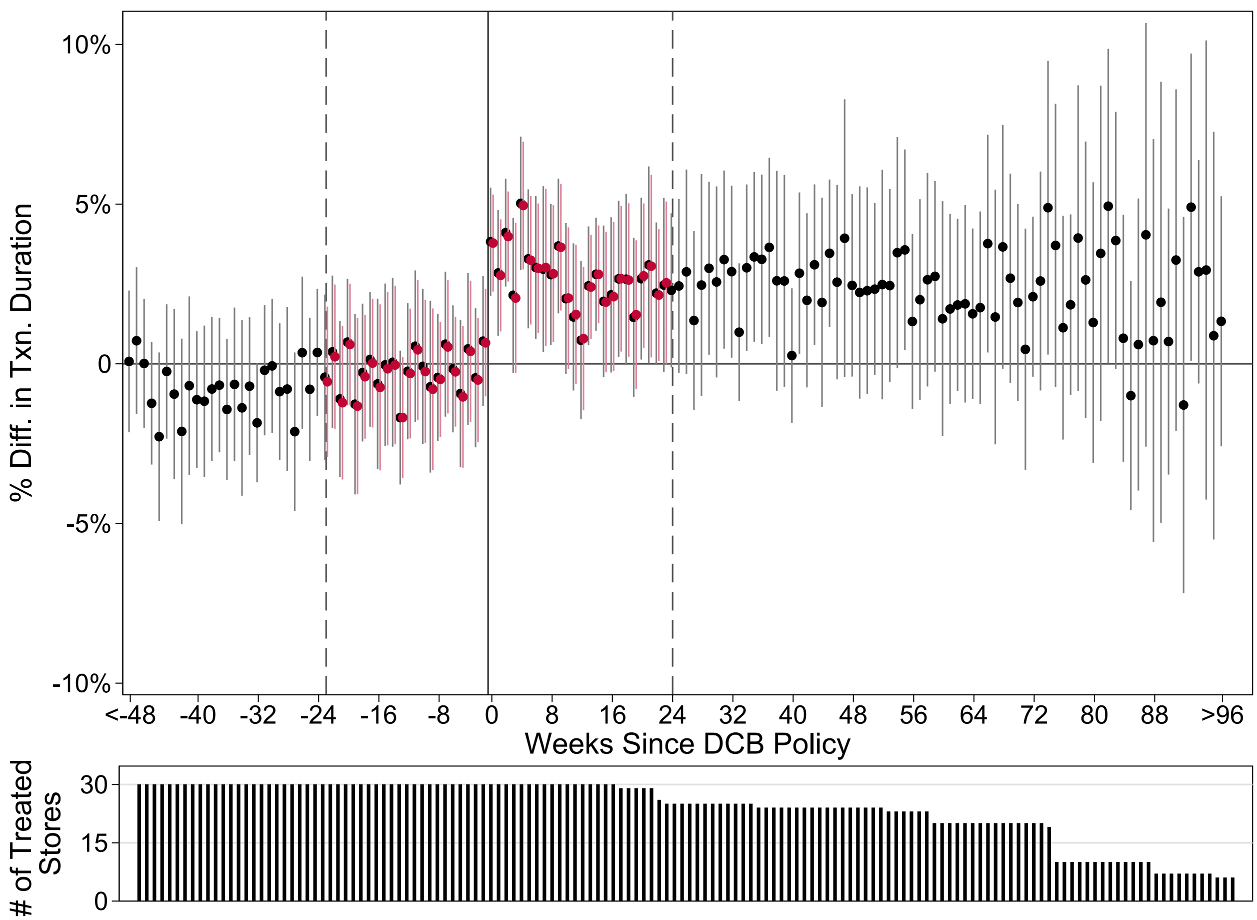


(b) Logged Transaction Duration—With Control Variables



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 90% 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—alcohol and tobacco, floral department items, fresh meat and seafood, fresh produce, pet items, and baby items.

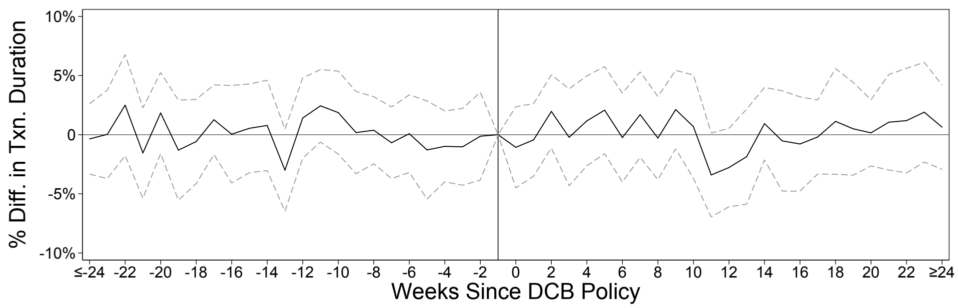
Figure 4: Effect of DCB Policies on Transaction Duration, with Extended Endpoints (*Store-Week Averages*)



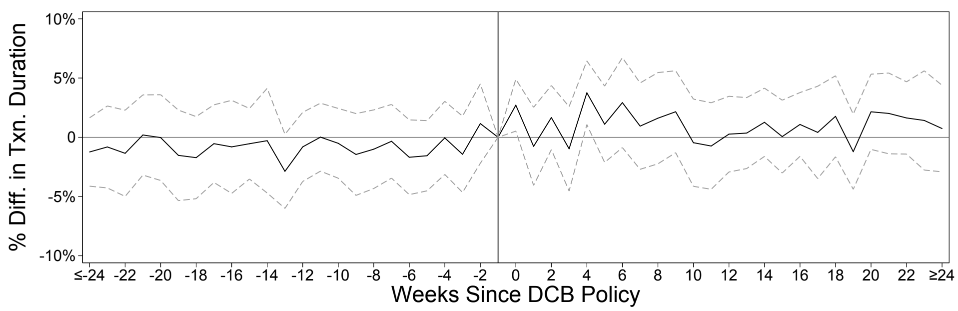
Note: The top panel of the figure displays the $\hat{\beta}_l$ event study estimates, in black, from the extended event study equation: $Y_{sjw} = \sum_{l=-49}^{97} \beta_l D_{l,jw} + \beta_x X_{sjw} + \theta_{sj} + \delta_w + \epsilon_{sjw}$. In red, between the dashed lines, are the β_l estimates from the full specification of Equation 1—same as in Figure 3b—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 90% 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 30 treated stores have $D_{15,sw} = 1$, only 6 treated stores have $D_{96,sw} = 1$.

Figure 5: Heterogeneity in Effect of DCB Policies on Transaction Duration, by Transaction Size Quartile (*Store-Week Averages*)

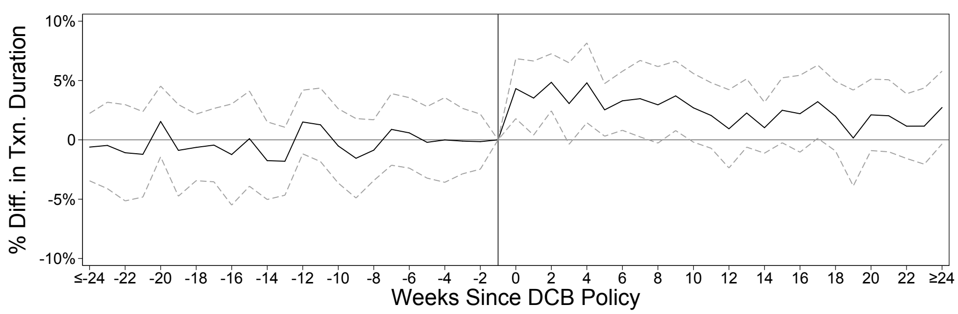
(a) Logged Transaction Duration—Q1 (Scans < 4)



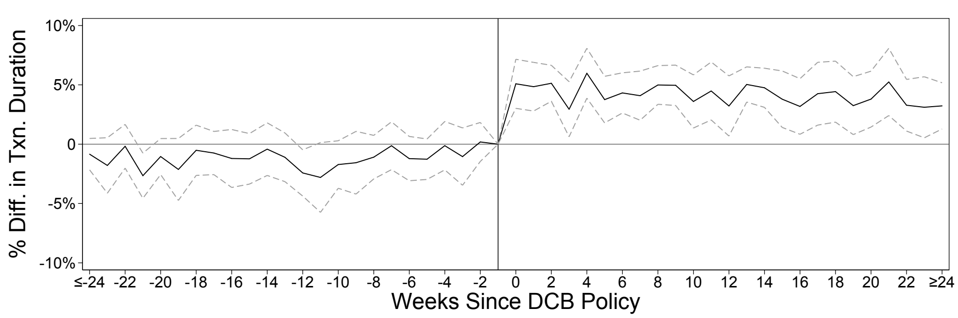
(b) Logged Transaction Duration—Q2 (Scans = 4-8)



(c) Logged Transaction Duration—Q3 (Scans = 9-18)



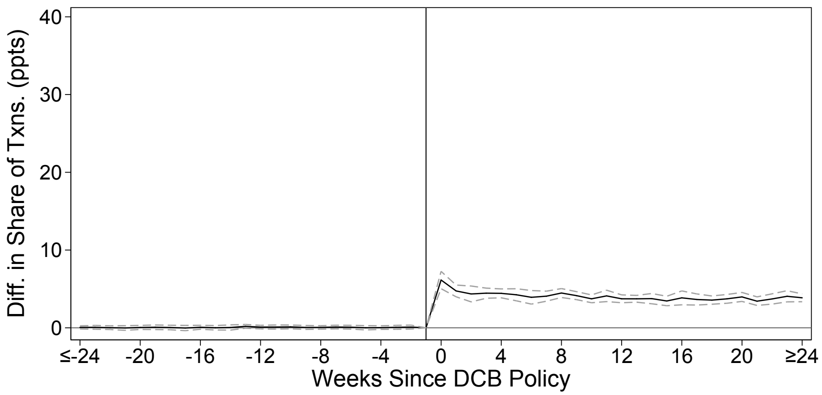
(d) Logged Transaction Duration—Q4 (Scans > 18)



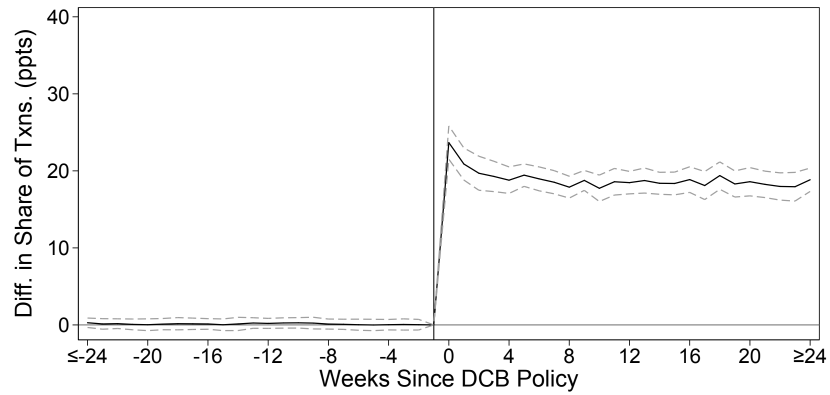
Note: The figure panels display the $\hat{\beta}_1$ coefficient estimates from full specification of event study Equation 1. The dependent variable is logged average transaction duration, measured in minutes, in store s, jurisdiction j, and week-of-sample w, by transaction size quartile with the smallest transaction quartile in panel (a) and the largest transaction quartile in panel (d). Upper and lower 90% confidence intervals are depicted in gray, estimated using two-way cluster robust standard errors on policy jurisdiction and week-of-sample.

Figure 6: Effect of DCB Policies on Share of Transactions Purchasing Paper Bags, by Transaction Size Quartile (*Store-Week Averages*)

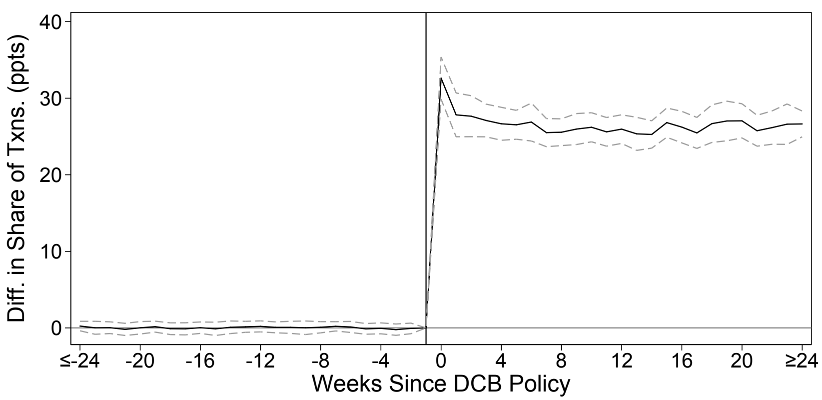
(a) Share Purchasing Paper—Q1 (Scans < 4)



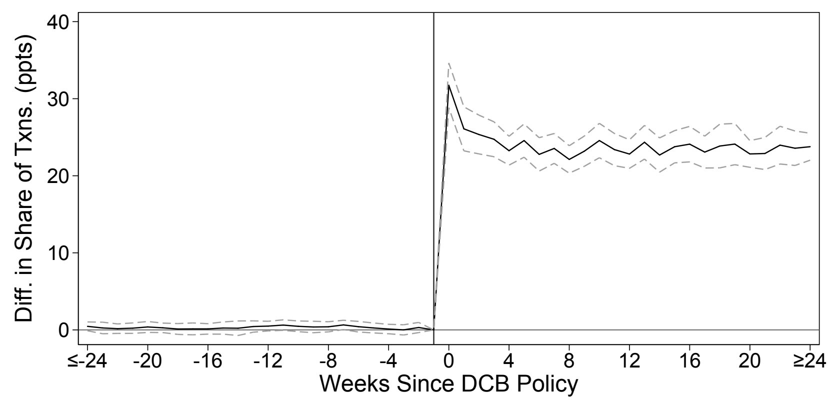
(b) Share Purchasing Paper—Q2 (Scans = 4-8)



(c) Share Purchasing Paper—Q3 (Scans = 9-18)



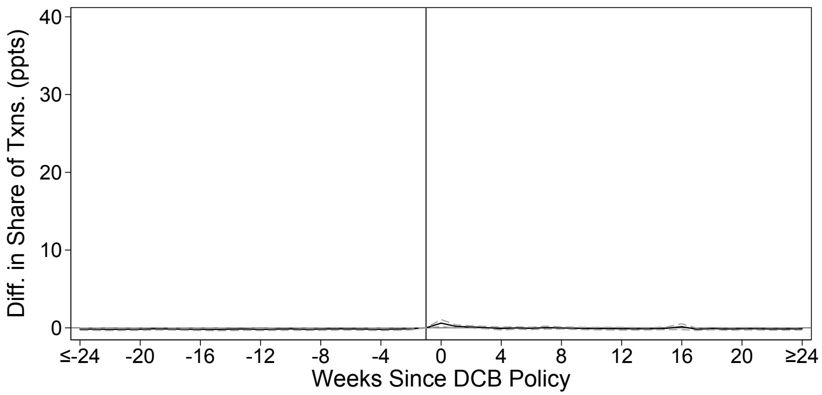
(d) Share Purchasing Paper—Q4 (Scans > 18)



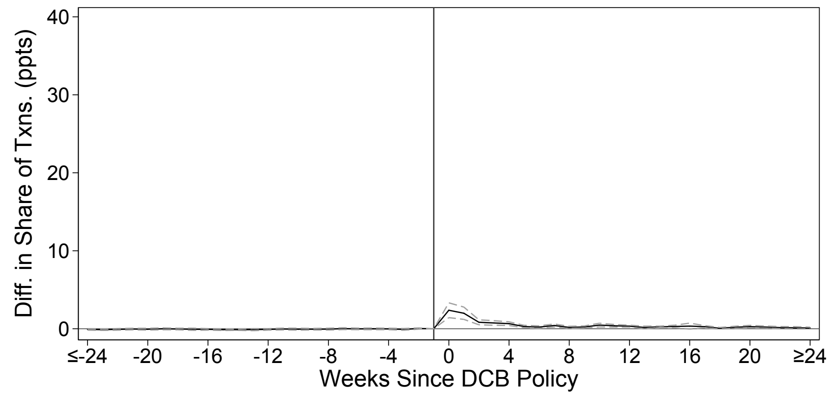
Note: The figure panels display the $\hat{\beta}_1$ coefficient estimates from the full specification of event study Equation 1. The dependent variable is share of transactions in store s, jurisdiction j, and week-of-sample w purchasing paper bags by transaction size quartile, with the smallest transaction quartile in panel (a) and the largest transaction quartile in panel (d). Upper and lower 90% 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 Share of Transactions Purchasing Reusable Bags, by Transaction Size Quartile (*Store-Week Averages*)

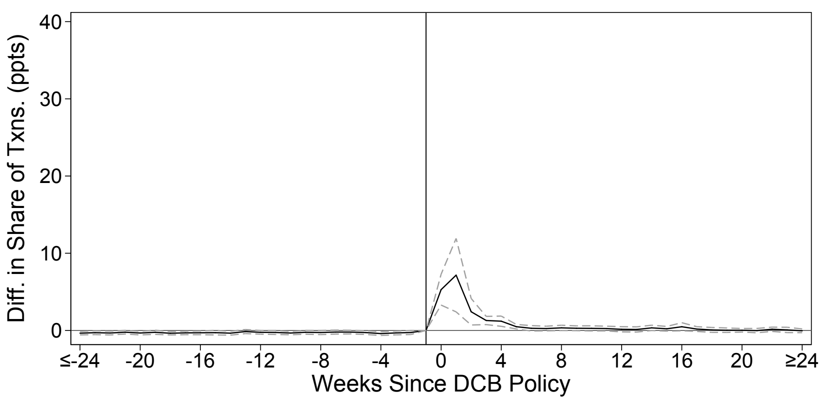
(a) Share Purchasing Reusable—Q1 (Scans < 4)



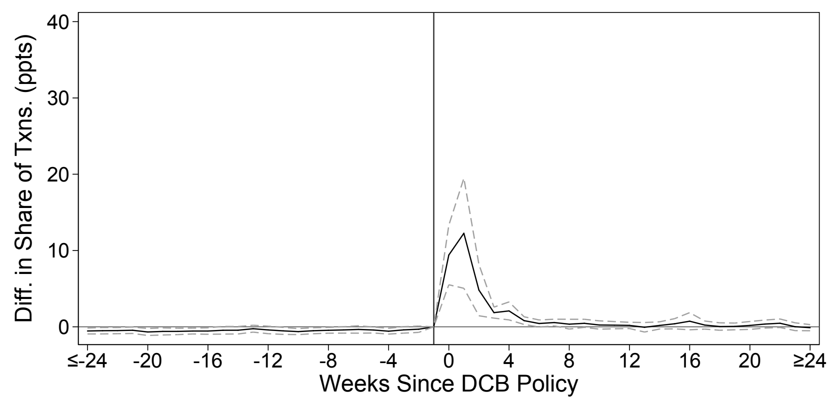
(b) Share Purchasing Reusable—Q2 (Scans = 4-8)



(c) Share Purchasing Reusable—Q3 (Scans = 9-18)



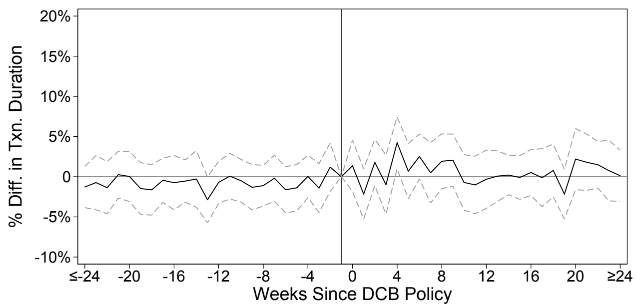
(d) Share Purchasing Reusable—Q4 (Scans > 18)



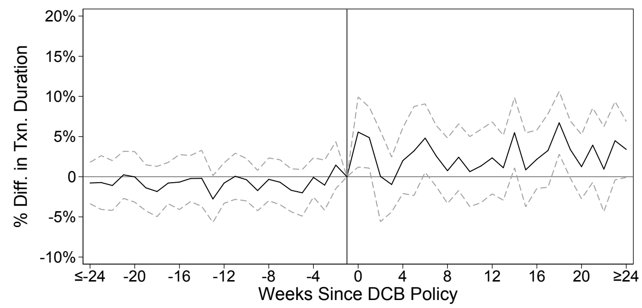
Note: The figure panels display the $\hat{\beta}_1$ coefficient estimates from the full specification of event study Equation 1. The dependent variable is the share of transactions in store s, jurisdiction j, and week-of-sample w purchasing reusable bags, by transaction size quartile with the smallest transaction quartile in panel (a) and the largest transaction quartile in panel (d). Upper and lower 90% confidence intervals are depicted in gray, estimated using two-way cluster robust standard errors on policy jurisdiction and week-of-sample.

Figure 8: Heterogeneity in Effect of DCB Policies on Transaction Duration, by Transaction Size Quartile and Paper Bag Purchase (*Store-Week Averages*)

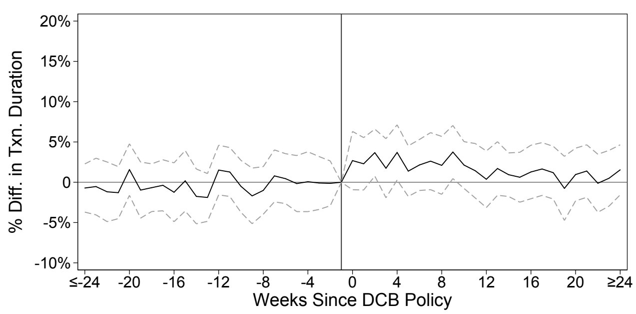
(a) Log Txn Duration—Q2
Without paper bags



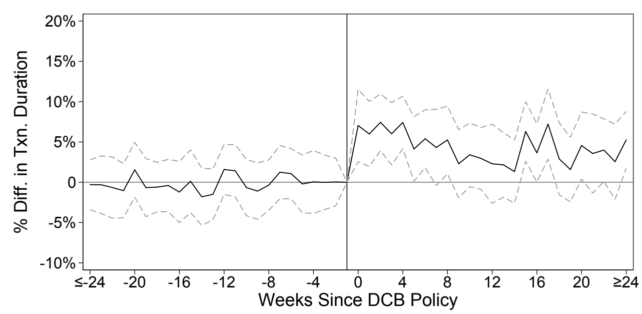
(b) Log Txn Duration—Q2
Paid paper bag fee



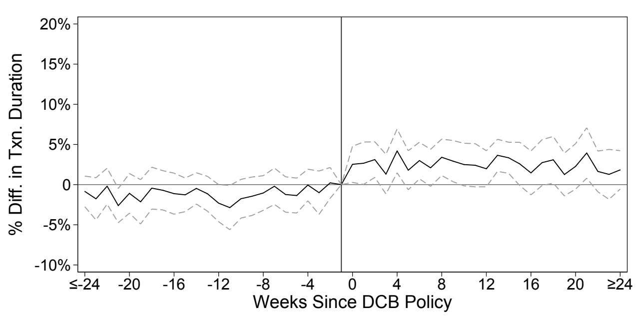
(c) Log Txn Duration—Q3
Without paper bags



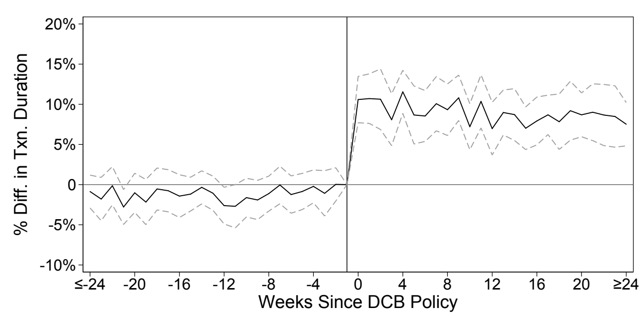
(d) Log Txn Duration—Q3
Paid paper bag fee



(e) Log Txn Duration—Q4
Without paper bags



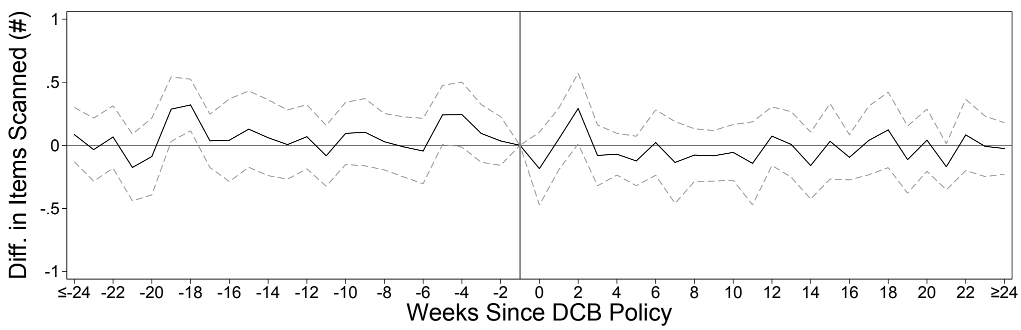
(f) Log Txn Duration—Q4
Paid paper bag fee



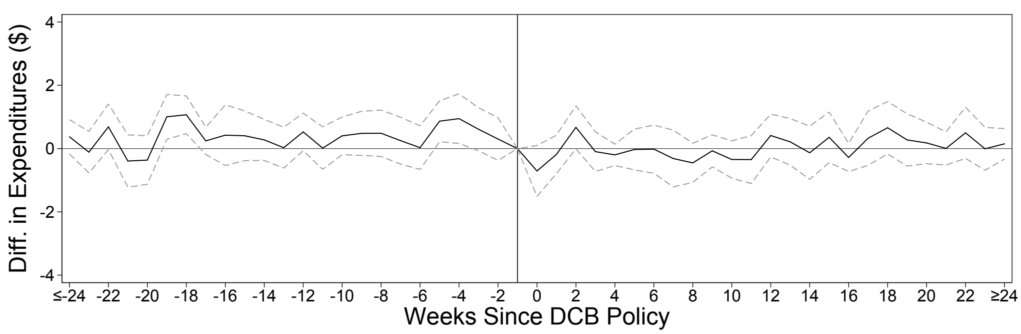
Note: The figure panels display the $\hat{\beta}_1$ coefficient estimates from the full specification of event study Equation 1. The dependent variable is logged average transaction duration, measured in minutes, in store s, jurisdiction j, and week-of-sample w, by transaction size quartile and paper bag use. Upper and lower 90% 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 Number of Items Scanned and Amount Spent (*Store-Week Averages*)

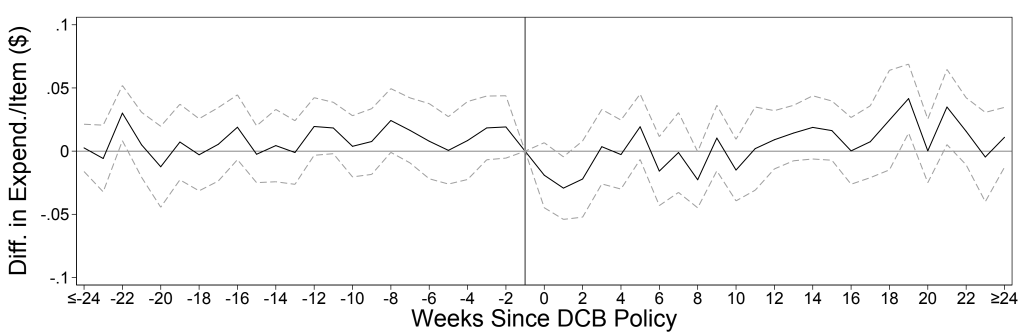
(a) Number of Items Scanned per Transaction, without Bags



(b) Amount Spent per Transaction ($)



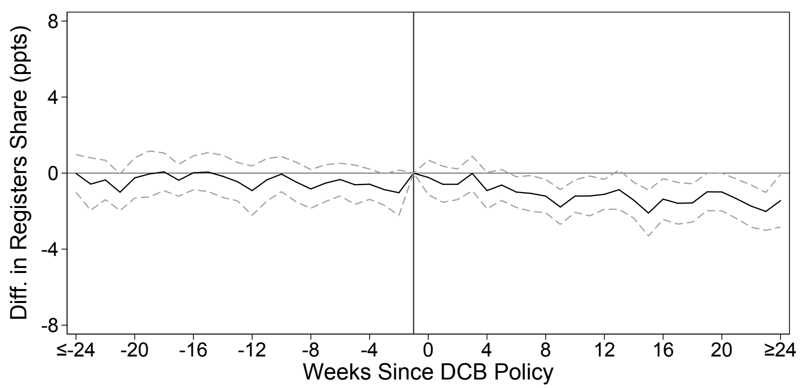
(c) Amount Spent per Item ($)



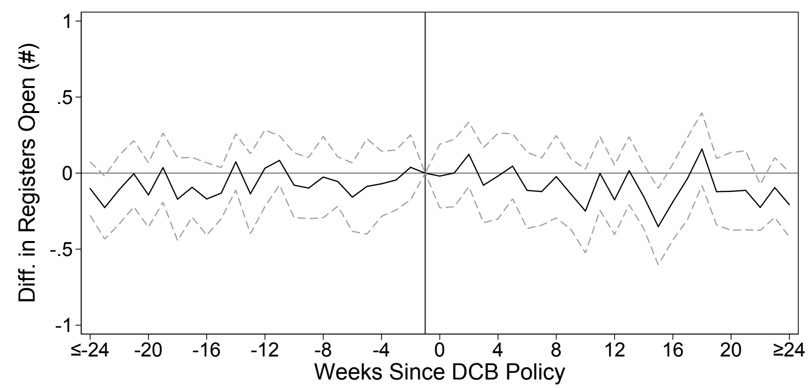
Note: The figures display the $\hat{\beta}_l$ coefficient estimates from the simple specification of event study Equation 1, controlling for only store and week-of-sample fixed effects. The dependent variables are (a) the average number of items scanned per transaction in store s, jurisdiction j, and week w, not including checkout bags, (b) the average amount spent per transaction in store s, jurisdiction j, and week w, and (c) the average amount spent per item in store s, jurisdiction j, and week w. Upper and lower 90% 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 Register Choice and Registers Open, by Register Type (*Store-Week Averages*)

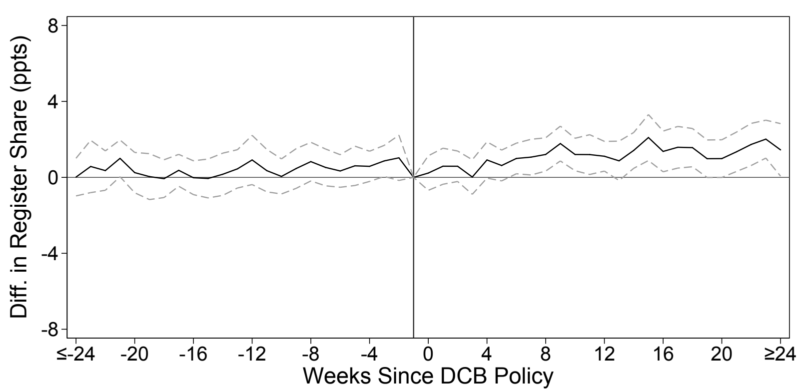
(a) Cashier-operated Register Share (full-service + express)



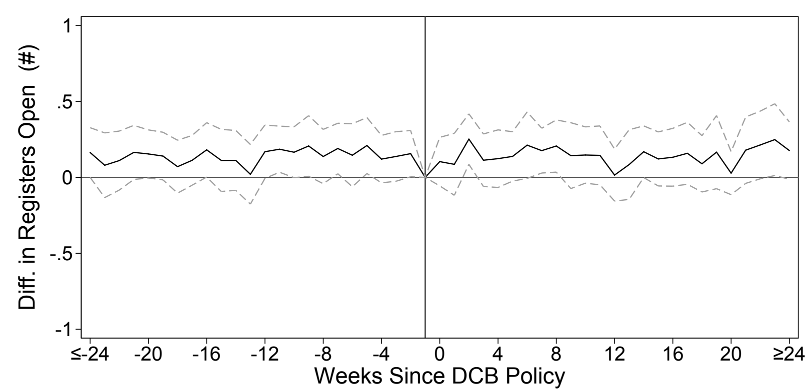
(b) Cashier-operated Registers Open (full-service + express)



(c) Self-checkout Register Share

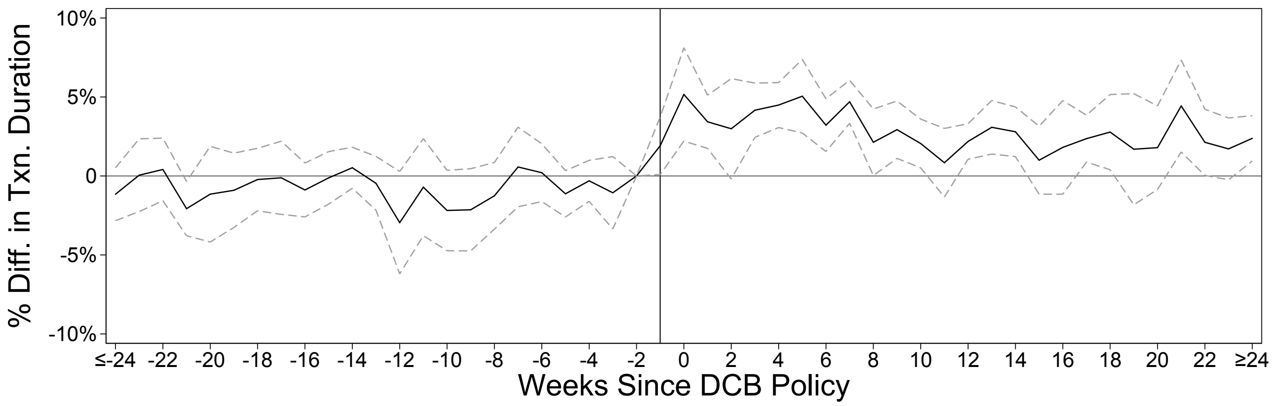


(d) Self-checkout Registers Open



Note: The figures display the $\hat{\beta}_l$ coefficient estimates from the simplest specification of Equation 1. The dependent variables in panels (a) and (c) are the share of transactions by register type in store s, jurisdiction j, and week-of-sample w. The dependent variables in panels (b) and (d) are the number of registers open by register type in store s, jurisdiction j, and week-of-sample w. Upper and lower 90% confidence intervals are depicted in gray, estimated using two-way cluster robust standard errors on policy jurisdiction and week-of-sample.

Figure 11: Effect of DCB Policies on Transaction Duration (*Cashier-Week Averages*)



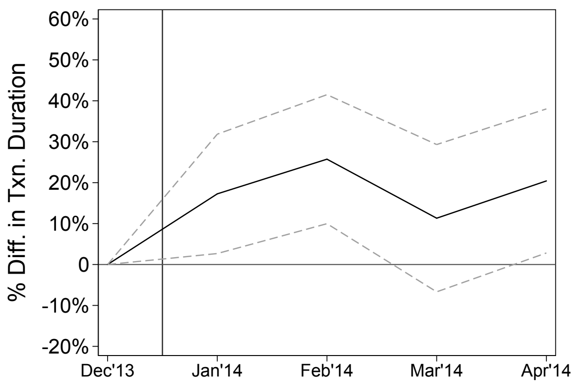
Note: Figure presents the full specification of event study Equation 3, with cashier and week-of-sample fixed effects and control variables for the average number of items scanned per transaction, average amount spent per transaction, the types of items purchased, and the weeks of experience of cashier 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 90% confidence intervals are depicted in gray, estimated using two-way cluster robust standard errors on policy jurisdiction and week-of-sample.

Figure 12: Learning Curve: Starting a New Shift vs. DCB Policies (*Cashier-Week Averages*)

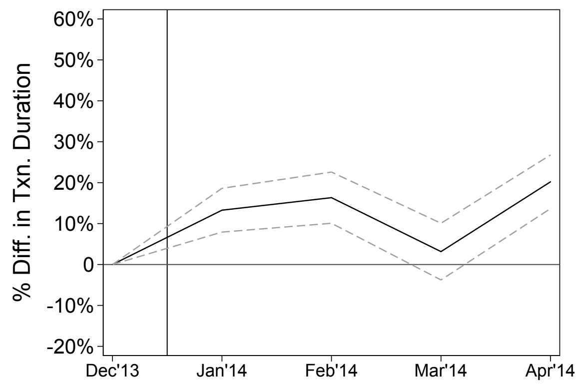
Note: Figure presents the results from equations 4 and 5, with cashier and week-of-sample fixed effects and control variables for the average number of items scanned per transaction, average amount spent per transaction, and the types of items purchased for cashier c in store s, jurisdiction j, and week-of-sample 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. The η_e estimates (for cashiers learning to work a new shift) are plotted on the left side of the graph. The post-policy $\hat{\beta}_l$ estimates (for cashiers learning after the policy change) are plotted on the right side of the graph. The estimates for cashiers at treated stores are depicted with red circles and the estimates for cashiers at the control stores are depicted with blue hollow circles. Upper and lower 90% confidence intervals are depicted in gray, estimated using two-way cluster robust standard errors on policy jurisdiction and week-of-sample.

Figure 13: Effect of DCB Policies on Transaction Duration and Bagger Presence (*In-store vs. Scanner Data*)

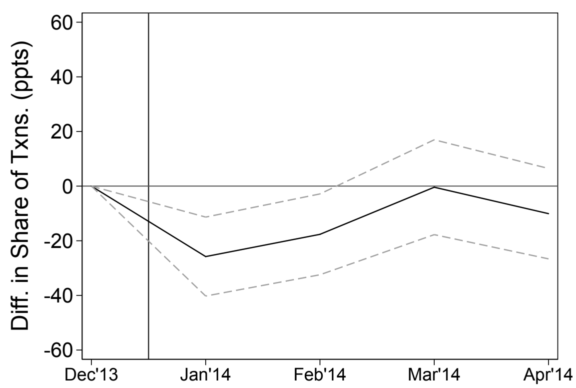
(a) Full-service Registers (*In-store Data*)



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



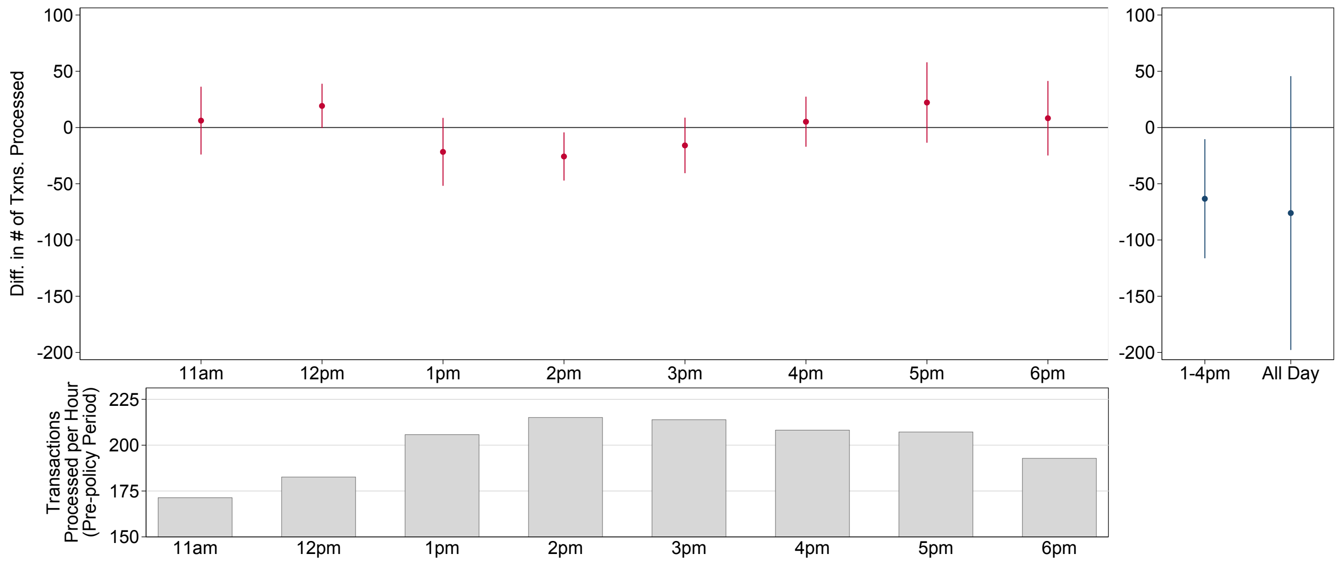
(c) Bagger Presence (*In-store Data*)



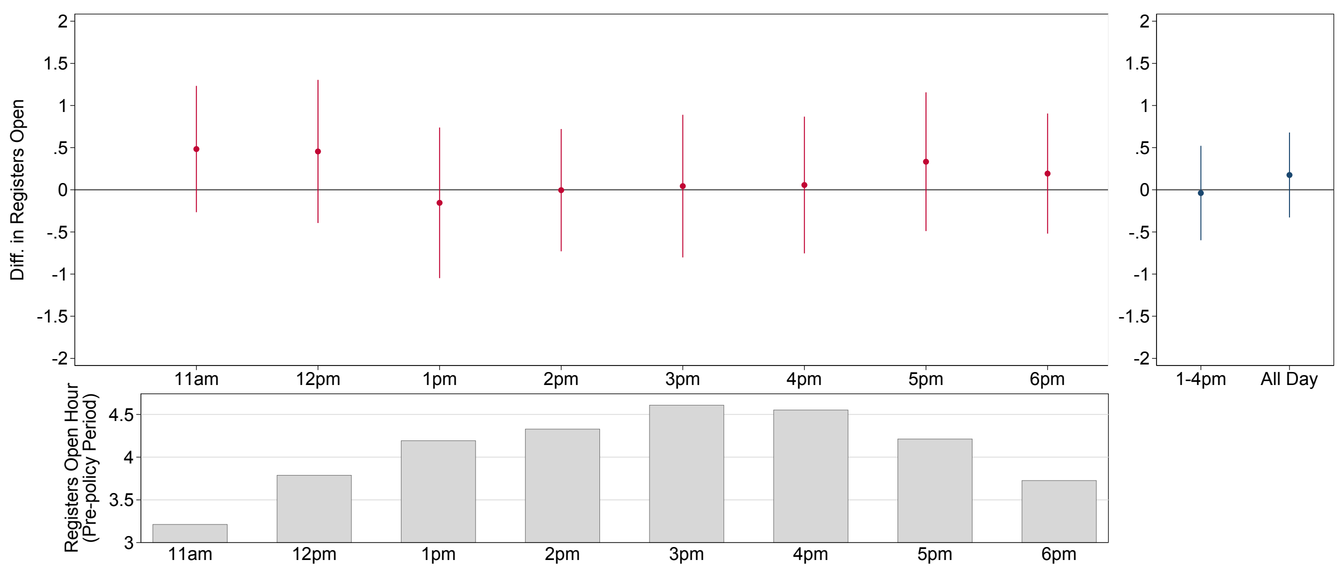
Note: The figure panels display the $\hat{\beta}_l$ coefficient estimates from the full specification of event study Equation 7. The dependent variable in panels (a) and (b) is logged average transaction duration, measured in minutes, of transaction t in store s and day-of-sample d. The dependent variable in panel (c) is an indicator equal to 1 if transaction t had a bagger present. Panels (a) and (c) use observational data collected in store while panel (b) uses scanner data. This analysis includes only transactions occurring at full-service registers, and not express or self-checkout registers. With the observational data in panel (a), 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 14: Effect of DCB Policies by Hour and by Shift (*Scanner Data from Three-Store Sample*)

(a) Number of Transactions Processed by Hour/Shift

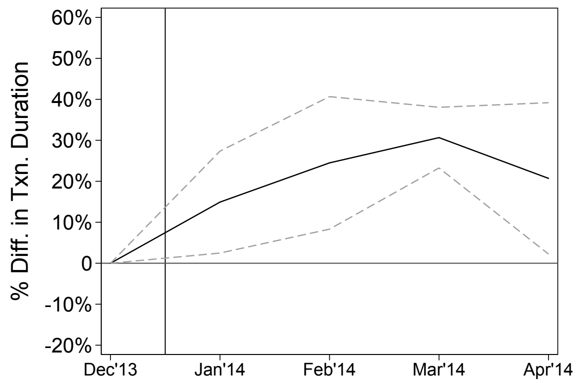


(b) Number of Registers Open by Hour/Shift



Note: The figure panels display the $\hat{\beta}_D$ coefficients estimates from the difference-in-differences Equation 2. The dependent variable in panel (a) is the number of transaction completed (by hour or by shift) in store s and week w. The dependent variable in panel (b) is the number of registers open (by hour or by shift) in store s and week w. Upper and lower 90% confidence intervals using robust standard errors are depicted in gray.

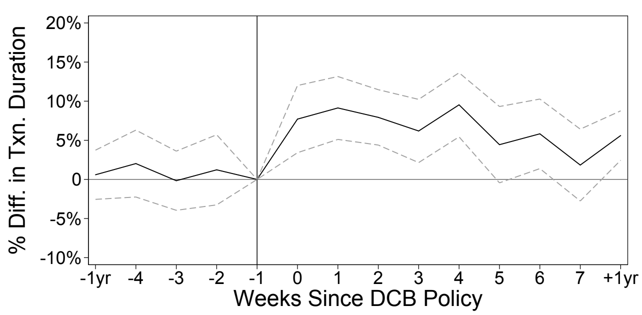
Figure 15: Effect of DCB Policies on Transaction Duration (*Discount Chain In-store Data*)



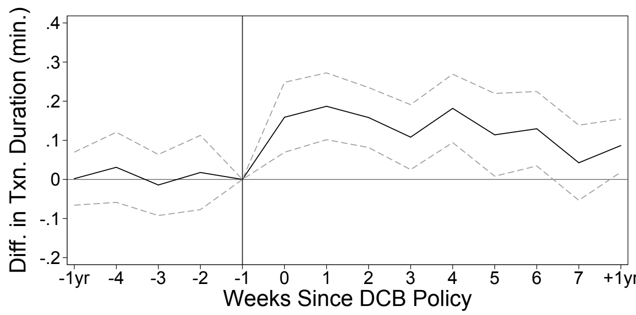
Note: The figure displays the $\hat{\beta}_l$ coefficient estimates from the full specification of event study Equation 7. The dependent variable is logged average transaction duration, measured in minutes, of transaction t in store s and day-of-sample d. This figure uses observational data collected in store at full-service registers. The control variables include indicators for the gender and race of the person paying, whether there was a checkout interruption, and the register number. Upper and lower 90% confidence intervals are depicted in gray. Standard errors are calculated using error clustering at the store-day level.

Figure 16: Heterogeneity in Effect of DCB Policies on Transaction Duration, by Transaction Size and DCB Purchase (*DC Data, Store-Week Averages*)

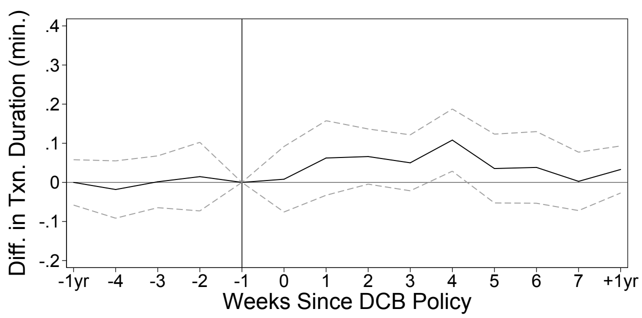
(a) Logged Transaction Duration—Full Sample



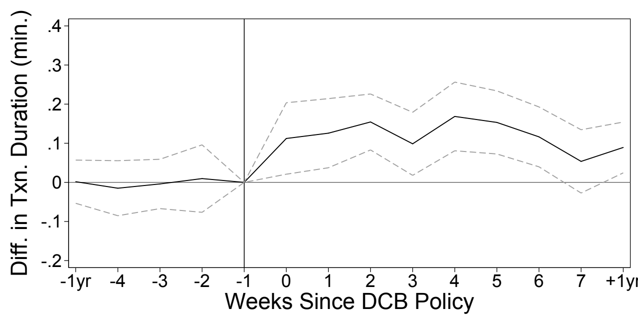
(b) Transaction Duration—Full Sample



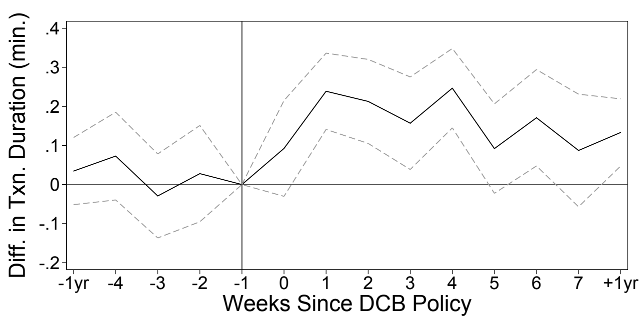
(c) Transaction Duration—Scans < 8
Without paying bag fee



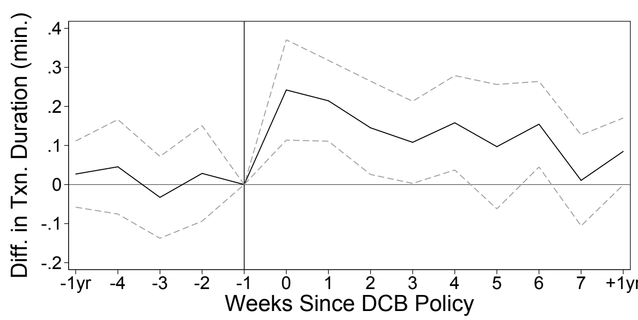
(d) Transaction Duration—Scans < 8
Paid bag fee



(e) Transaction Duration—Scans ≥ 8
Without paying bag fee



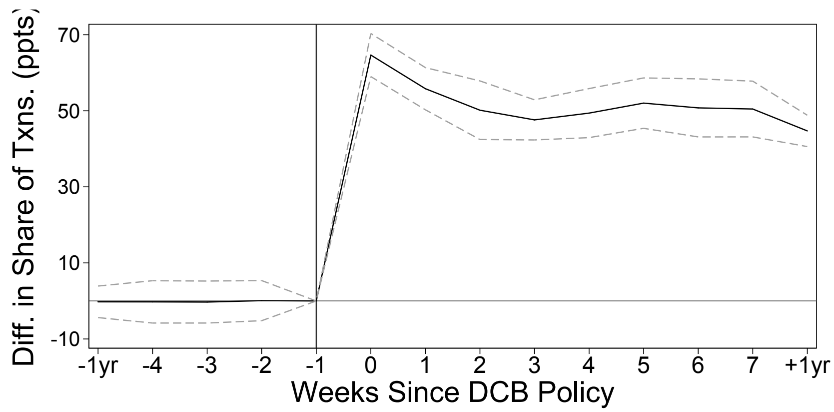
(f) Transaction Duration—Scans ≥ 8
Paid bag fee



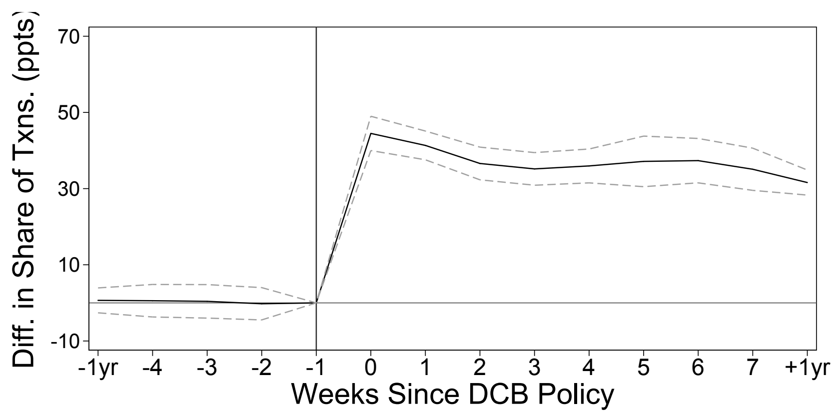
Note: The figures display the $\hat{\beta}_l$ coefficient estimates from the full specification of event study Equation 8. The dependent variable is average transaction duration, measured in minutes, in store s and week-of-sample w. The dependent variable is in logs in panel (a) and in levels in all other panels. Upper and lower 90% confidence intervals using robust standard errors are depicted in gray. 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).

Figure 17: Effect of DCB Policies on Share of Transactions Paying Bag Fee, by Transaction Size (*DC Data, Store-Week Averages*)

(a) Share Paying Fee—Scans ≥ 8



(b) Share Paying Fee—Scans < 8



Note: The figures display the $\hat{\beta}_l$ coefficient estimates from the full specification of event study Equation 8. The dependent variable is the share of transactions paying the bag fee. Upper and lower 90% confidence intervals using robust standard errors are depicted in gray. 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).

10 Tables

Table 1: Average Store Characteristics and Demographics

	(1) Control Stores	(2) Treat Stores	(3) P-value of Diff.	(4) California	(5) United States
Store Characteristics					
Building Size (ft^2)	42,008.47	44,320.27	0.520		
Open Year	1985	1984	0.751		
Last Remodel Year	2005	2005	0.733		
Departments & Services (share)					
Bakery	0.79	0.80	0.931		
Pharmacy	0.47	0.67	0.188		
Deli	1.00	0.97	0.432		
Floral	0.95	0.93	0.846		
Coffee Bar	0.74	0.53	0.161		
Gas Station	0.05	0.07	0.846		
Juice Bar	0.11	0.10	0.954		
Sandwich Counter	0.05	0.07	0.846		
Self-checkout registers (share)	0.47	0.53	0.692		
Store Location Demographics					
Median Income ($)	$64,025	$63,027	0.857	$47,493	$41,994
Household Size (#)	2.59	2.53	0.506	2.87	2.59
White (share)	0.71	0.69	0.709	0.60	0.75
Black (share)	0.05	0.05	0.977	0.07	0.12
Asian (share)	0.11	0.11	0.759	0.11	0.04
Over 65 (share)	0.12	0.12	0.808	0.11	0.12
Do not own vehicle (share)	0.06	0.07	0.316	0.10	0.10
Urban (share)	0.74	0.87	0.262	0.87	0.79
N Stores	19	30			

Note: Asterisks indicate the following: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *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-Level Characteristics in 2011

	Control	Treat	P-value of Diff.
Transaction Duration (minutes)	1.99	2.13	0.364
Full-Service	1.91	2.01	0.140
Express	1.38	1.49	0.024**
Self-Checkout	3.70	3.98	0.078*
Items Scanned (#)	13.32	13.54	0.699
Full-Service	19.09	19.58	0.695
Express	7.64	8.65	0.035**
Self-Checkout	6.45	6.62	0.473
Amount Paid ($)	39.52	40.18	0.728
Full-Service	56.12	57.70	0.684
Express	23.35	26.15	0.070*
Self-Checkout	19.64	19.87	0.750
N Stores	19	30	–
N Stores w/Self-checkout	9	16	–

Note: 2011 is in the pre-policy period for all stores in the sample. Asterisks indicate the following: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Source*: Author's calculations from the scanner data.

Table 3: Average Store-Shift Characteristics in 2011

	Control	Treat	P-value on Diff.
Stores with Self-Checkout			
Transactions per shift (#)	689.91	704.75	0.755
Full-Service (share)	0.53	0.50	0.274
Express (share)	0.17	0.21	0.292
Self-Checkout (share)	0.32	0.30	0.378
Amount spent per shift ($)	$28,623.03	$29,086.14	0.866
Full-Service (share)	0.77	0.73	0.113
Express (share)	0.08	0.13	0.052*
Self-Checkout (share)	0.15	0.14	0.572
Items bought per shift (#)	9,566.03	9,811.68	0.783
Full-Service (share)	0.77	0.73	0.099*
Express (share)	0.08	0.13	0.047**
Self-Checkout (share)	0.15	0.14	0.603
Registers open per shift (#)	11.29	11.80	0.445
Register Capacity (#)	13.33	14.06	0.233
N Stores	9	16	–
Stores without Self-Checkout			
Transactions per shift (#)	438.94	446.95	0.809
Full-Service (share)	0.55	0.52	0.327
Express (share)	0.45	0.49	0.311
Amount spent per shift ($)	$14,822.13	$15,373.56	0.755
Full-Service (share)	0.73	0.68	0.144
Express (share)	0.27	0.32	0.144
Items bought per shift (#)	5,045.90	5,181.37	0.809
Full-Service (share)	0.74	0.69	0.126
Express (share)	0.26	0.31	0.126
Registers open per shift (#)	5.29	5.53	0.562
Register Capacity (#)	7.70	7.43	0.533
N Stores	10	14	–

Note: 2011 is in the pre-policy period for all stores in the sample. Asterisks indicate the following: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Source*: Author's calculations from the scanner data.

Table 4: Effect of DCB Policies on Transactions per Shift and Line Length
(*Store-Week Averages*)

	(1) Txns. per Shift (#)	(2) Txn. Duration (mins)
Levels (Y_{sjw})		
Ban Effective Dummy	-19.261***	0.060***
	(6.519)	(0.019)
Percent (lnY_{sjw})		
Ban Effective Dummy	-0.032***	0.030***
	(0.010)	(0.008)
Num of Obs.	8673	8673
Standard Errors	Cluster	Cluster
Covariates X_{sjw}	Yes	Yes
Store FE	Yes	Yes
Week-of-sample FE	Yes	Yes
Mean Y_{sw} (2011)	573.142	2.079
Changes in Line Length		
Ave. ↑ customers in line ($\frac{\lvert\hat{\beta}_D\rvert}{2}$)	9.631	
Registers open, post-policy	9.276	
Ave. ↑ customers in line, per open register	1.038	
Registers available	10.735	
Ave. ↑ customers in line, per available register	0.897	

Note: The top half of the table presents the results from Equation 2, with the outcome variable estimate in both levels and logs. In column (1), the outcome variable is the average number of transactions completed per 3 hour shift in store s, jurisdiction j, and week w. In column (2), the outcome variable is the average transaction duration, measured in minutes, in store s, jurisdiction j, and week w. Standard errors are in parentheses. Standard errors are estimated using two-way error clustering at the policy jurisdiction and week-of-sample level. Asterisks indicate the following: $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

The bottom half of the table reports the $\frac{|\hat{\beta}_D|}{2}$ estimate from the specification in levels. This estimate is the average increase in customers in line due to the policy change and it is used to calculate the additional number of customers in line *per register* either (i) given the average number of registers open in the post-policy period, or (ii) if all existing registers were open.

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	(1.114)	(1.234)	
N	1,692	34,028	
Stores with DCB Policy	1.756	1.910	-0.154***
SD	(1.144)	(1.252)	
N	934	17,562	
Stores without DCB Policy	1.670	1.759	-0.089**
SD	(1.073)	(1.210)	
N	758	16,466	

Note: Standard deviations in parentheses. Asterisks indicate the following: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *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.

A Appendix

A.1 Customer Level Sensitivity Analysis

In this appendix, I examine whether the main results are sensitive to the inclusion of customer fixed effects, using scanner data averaged to the customer-month frequency. The event study model at the customer-month level is as follows:

$$Y_{isjm} = \sum_{l=-8}^{6} \beta_l D_{l,jm} + \beta_x X_{isjm} + \gamma_{isj} + \chi_m + \epsilon_{isjm} \tag{9}$$

The outcome variable, Y_{isjm}, is the logged average transaction duration for customer i in store s, jurisdiction j, and month-of-sample m. The control variables include the average number of items scanned, the average amount spent, the average types of items purchased by customer i in store s and month m, as well as month-of-sample and customer fixed effects. Additionally, I control for the number of months the customer has appears in the sample. Including customer fixed effects, γ_{isj}, means the β_l coefficients in Equation 9 measure the policy effects *within customers* over time.

It is important to note that I designed the data to capture stores and cashiers over time, and not necessarily customers over time. I have every transaction during the 1:00-4:00pm shift on Saturdays and Sundays for 49 stores and 3.5 years. If a customer happens to shop at the same time each week or month *using their rewards card*, I can see them multiple times in the sample using the identification code of their card. If they do not use a customer card or they shop at different times each week, I cannot track the customer.

There are 2,033,434 unique customer codes in my sample. 48% of customers appear only once, 24% of customers appear 2-3 times, 12% appear 4-6 times, and the remaining 16% appear 7 times or more. I drop all customers that appear more than 100 times, which is less than half a percentage point of customers. I do this because high card frequencies may be driven by cashiers scanning a store-owned rewards card, or their own personal rewards card, when customers do not have a card. I also drop the customers that appear less than twelve times in the sample, in order to focus on customers that are in the sample long enough to experience learning. As such, the results I present show how DCB policies affect checkout speed within customers, for the 137,031 customers that shop regularly.

Appendix Figure A.1 presents the event study results from estimating Equation 9. The coefficients found in panel (a) at the customer-level are similar to those at the cashier-level in Figure 11. During the first month of the policy, transactions are 5.8% longer ($\hat{\beta}_0 = 0.058$). Transactions that occur 6 months or more after the policy implementation remain 3.9% longer ($\hat{\beta}_6 = 0.039$). Using Wald tests to compare the coefficients, I cannot reject that all $\hat{\beta}_l$ coefficients in the post-policy period are the same as one another nor can I reject that $\hat{\beta}_0 = \hat{\beta}_6$. Therefore, I cannot conclude that the customers who shop regularly become significantly faster at checking out over time after the policies.

I next split the customers at treated stores into four groups by whether they *ever* buy paper bags and by whether they buy fewer than 8 items on average.[84] Panels (b) to (e) of Appendix Figure A.1 present the results of Equation 9 estimated separately for each of these treated groups. Importantly, the control customers are the same in each panel. I find no statistically significant slowdown due to the DCB policies for the smaller transaction customers, both for those that never purchase paper (panel b) and for those that ever purchase paper (panel c).[85] Conversely, I find a persistent 4.4% slowdown in transaction duration for the larger transaction customers that never purchase paper (panel d), and an 8.5% slowdown for the larger transaction customers that ever purchase paper (panel e). These heterogeneity results match what I found when using the store-week data in Figure 8. This is reassuring because, at the store-week level, splitting the transactions by whether a paper bag was purchased in the post-policy period meant that the treated customers in the pre-period were not necessarily the same as the treated customers in the post-period. However, the near-zero pre-policy $\hat{\beta}_l$ coefficients in every panel in Appendix Figure A.1 indicates that before the policy, none of the treated customers differed in transaction duration from the control customers. Yet after the policy, treated customers with larger transactions that choose paper bags have significantly longer transaction durations than control customers.

I also estimate Equation 9 with the following outcome variables: (1) average number of items

[84] 13% of customers in treated stores purchase smaller transactions and never purchase paper bags, 10% purchase smaller transactions and purchase a paper bag at least once in the sample, 34% purchase larger transactions and never purchase paper bags, and the remaining 43% purchase larger transactions and purchase a paper bag at least once in the sample.

[85] While not statistically significant, the DCB policies lead to a temporary slowdown for smaller transaction customers purchasing paper bags, with a $\hat{\beta}_0 = 0.039$ in panel (c).

bought per transaction, (2) average amount spent per transaction, (3) average amount spent per item, (4) share of transactions completed at self-checkout registers, and (5) the number of transactions completed per month. In none of these regressions do I find significant changes in the outcome variable that is contemporaneous with the policy change. Thus for the sample of customers that shop regularly, the DCB policies do not appear to be altering the amount of items they buy or the amount they frequent a store during the 1:00-4:00pm weekend shift.[86]

A.2 Correlates of Customer Paper Bag Choice

To understand what factors correlate with choosing to pay the paper bag fee, I estimate the following model using post-policy customer level data at treated stores:

$$Paper_{isjm} = \beta_x X_{isjm} + \sum_{l=0}^{24} \beta_l D_{l,jm} + \gamma_{isj} + \chi_m + \epsilon_{isjm} \tag{10}$$

where $Paper_{isjm}$ is the share of customer i's transactions in store s, jurisdiction j, and month m where at least one paper bag is purchased. X_{isjm} is a set of customer level and store level covariates, $D_{l,jm}$ is a set of months-since-policy dummy variables, γ_{isj} are customer fixed effects, and χ_m are month-of-sample fixed effects. Appendix Table A.3 presents the results, with only month-of-sample fixed effects and months-since-policy fixed effects in column (1) and additionally customer fixed effects in column (2).

The main takeaway from Appendix Table A.3 is that paper bag use in the post-policy period is positively correlated with income, transaction size, and the purchase of more expensive items. Similar to what was shown in Figure 6, I find that paper bag use increases with transaction size, except for the largest transactions in Q4, which are less likely to choose paper than Q3. Also, as the amount spent per item increases, paper bag use increases. Both with and without customer fixed effects, I find that paper bag use is negatively correlated with purchasing floral items and positively correlated with bakery and deli, prepared foods, meat and seafood, and shelf-stable food items. The type of register is not significantly correlated with paper bag use. The more trips a customer makes to the store per month (during the 1:00-4:00pm shift), the

[86] While not statistically different from zero, there is evidence that customers shop fewer times per month during the 1:00-4:00pm weekend shift in the post-policy period. This is consistent with customers choosing to shop at different times of the day or at different stores, or customers remaining in line longer and spilling over into the next shift.

less likely they are to use paper.

At the store level, the customers in stores with the highest median incomes (>$72K) are the most likely to pay for paper bags, followed by customers at stores with middle median incomes ($55-72K). Customers at store with the lowest median incomes (< $55K) are the least likely to purchase paper. Customers at stores in areas with larger Asian populations are the less likely to choose paper bags. Finally, paper bag use is positively correlated with a higher share of people not owning a vehicle and negatively correlated with the size of the store.

A.3 In-store Data Description

The in-store data were obtained through direct observation of transactions by enumerators stationed near checkout lanes. For each transaction, we collected data on the number and types of bags used, whether a bagger was present, the length of the transaction in minutes, and basic demographic data such as gender and race of the person paying. This type of transaction specific information can only be gained from in-store observations, and is not included in the scanner datasets from these stores. Four visits per store occurred in December 2013, before the Richmond DCB policy went into effect, and 4–6 visits occurred in January and February 2014, after the policy was in place. We also made an additional four visits in March and April 2014 to collect follow-up data. Each visit lasted 1 to 2 hours and was made on either Saturday or Sunday between 11:00am and 7:00pm.

We visited a total of seven stores, belonging to two different categories of grocery chains within the same treated and control cities. The first chain is the same supermarket chain as in the main analysis of this paper, for which I also have scanner data. It is a large national chain, offering high and low prices in many products. The other chain is a regional discount chain, offering name-brand products at closeout prices.

Appendix Table A.5 presents the pre-policy summary statistics from 2013, for the in-store data collected at the national chain supermarkets (columns 1–3) and at the discount chain supermarkets (columns 4–6). Transaction level averages are presented separately for stores with and without DCB policies. Since transactions in 2013 occur before the policy change, I group the treated and no-policy stores together in the *No DCB Policy* columns and the prior-policy stores in *DCB Policy* columns. The variables recorded at checkout include indicators for the presence of baggers, a transaction being interrupted,[87] and the gender and race of the customer paying. In addition, variables for the number and types of bags used were recorded.

For the national chain stores in columns 1–3, I find that the presence of baggers, checkout interruptions, and the gender of the person paying does not differ significantly between stores with and without policies. However, the stores without DCB policies have a higher share of Black customers than the stores with DCB policies. With respect to the number and types of bags used, bag usage differs greatly between stores with and without DCB policies. In particular, the share of customers using no bags, using paper bags, and bringing reusable bags is significantly higher in stores with DCB policies in place, while the share using thin plastic is zero. Comparing discount chain averages (columns 4–6) to the national chain averages (columns 1–3), I find that discount chain transactions are less likely to have baggers present and have a shorter transaction duration than those at the national chain. Discount chain customers are also less likely to be White and are more likely to use no bags when a DCB policy is in effect.

For a more detailed discussion of the in-store data and the effects of DCB policies on bag usage at checkout, please see Taylor and Villas-Boas (2016b).

A.4 Difference-in-Differences Estimation Using In-Store Data

In Section 6.1.1, I matched the transactions collected in-store at the national chain to their corresponding transactions in the scanner data. With these matched data, as well as the full scanner and in-store datasets for the same stores and days, I estimate the following difference-in-differences (DID) model:

$$Y_{tsjd} = \beta_D D_{jd} + \beta_x X_{tsjd} + \theta_{sj} + \delta_d + \epsilon_{tsjd} \tag{11}$$

where Y_{tsjd} is the logged transaction duration of transaction t in store s, jurisdiction j, on day-of-sample d, D_{jd} is an indicator for transactions at the treated store during the DCB policy effective period, X_{tsjd} is a set of control variables, θ_{sj} are store fixed effects and δ_d are day-of-sample fixed effects. Appendix Table A.6 reports the results. In column (1), I estimate Equation 11 with the scanner data, controlling for the number of items purchased and amount

[87] Interruptions include price checks, cashiers switching registers, phone calls, and the customer having payment issues such as card denial or paying by check

spent. I find that $\hat{\beta}_D = 0.104$, which means that DCB policies correspond with a 10.4% increase in transaction duration at the treated store relative to the no-policy and prior-policy stores. In column (3), I estimate the model with the in-store data, controlling for the presence of baggers, checkout interruptions, the gender and race of the person paying, and the number and types of bags used. I find that $\hat{\beta}_D = 0.155$. In column (2), I estimate the model with the matched data, controlling for the covariates from both the scanner and in-store data. Once again I estimate that DCB policies lead to a positive slowdown in transaction duration ($\hat{\beta}_D = 0.099$), however, the $\hat{\beta}_D$ estimate is no longer statistically significant when using the smaller matched dataset. Yet reassuringly, comparing columns (1) and (2), the lack of in-store covariates in column (1) does not bias its $\hat{\beta}_D$ estimate.

I also find interesting patterns with respect to the control variables themselves. Controlling for all else, each additional item scanned increases transaction duration by 1-2%, while each additional dollar spent increases transaction duration by 0.2%. The presence of baggers is associated with faster transaction duration, while experiencing a checkout interruption or payment issue is associated with much slower transaction durations. Men checkout 8-11% faster than women, while the race of the person paying is not correlated with differences in checkout speed.

I also estimate Equation 11 using the in-store data from the discount chain (column 4). Comparing in-store DID results in columns (3) and (4), both chains experience similar slowdowns due to the DCB policy. Transactions at the national chain are 15.5% slower due to the policy change, which equates to 0.224 minutes longer, and transactions at the discount chain are 33.5% slower, which equates to 0.178 minutes longer. Unlike at the national chain, the presence of a bagger does not significantly alter checkout duration, yet similar to the national chain, male shoppers are faster at checkout than female shoppers, all else equal.

A.5 Appendix Tables and Figures

Table A.1: List of California DCB Policies: 2007–2014

	Jurisdiction	County	Effective Date	Population (2010)	Population w/Ban (%)
1	San Francisco	San Francisco	Oct-07	805,235	2.2%
2	Malibu	Los Angeles	Dec-08	12,645	2.2%
3	Fairfax	Marin	May-09	7,441	2.2%
4	Palo Alto	Santa Clara	Sep-09	64,403	2.4%
5	Calabasas	Los Angeles	Jul-11	23,058	2.5%
6	Unincorporated Areas	Los Angeles	Jul-11	981,861	5.1%
7	Long Beach	Los Angeles	Aug-11	462,257	6.3%
8	Santa Monica	Los Angeles	Sep-11	89,736	6.6%
9	Unincorporated Areas	Marin	Jan-12	18,451	6.6%
10	San Jose	Santa Clara	Jan-12	945,942	9.2%
11	Unincorporated Areas	Santa Clara	Jan-12	97,882	9.4%
12	Unincorporated Areas	Santa Cruz	Mar-12	130,666	9.8%
13	Manhattan Beach	Los Angeles	Apr-12	35,135	9.9%
14	Monterey	Monterey	Jun-12	27,810	9.9%
15	Sunnyvale	Santa Clara	Jun-12	140,081	10.3%
16	Pasadena	Los Angeles	Jul-12	137,122	10.7%
17	Ojai	Ventura	Jul-12	7,461	10.7%
18	Carpinteria	Santa Barbara	Jul-12	13,040	10.7%
19	Solana Beach	San Diego	Aug-12	12,867	10.8%
20	Millbrae	San Mateo	Sep-12	21,532	10.8%
21	Watsonville	Santa Cruz	Sep-12	51,199	11.0%
22	Arroyo Grande	San Luis Obispo	Oct-12	17,252	11.0%
23	Atascadero	San Luis Obispo	Oct-12	28,310	11.1%
24	Grover Beach	San Luis Obispo	Oct-12	13,156	11.1%
25	Morro Bay	San Luis Obispo	Oct-12	10,234	11.2%
26	Paso Robles	San Luis Obispo	Oct-12	29,793	11.2%
27	Pismo Beach	San Luis Obispo	Oct-12	7,655	11.3%
28	San Luis Obispo	San Luis Obispo	Oct-12	45,119	11.4%
29	Unincorporated areas	San Luis Obispo	Oct-12	118,486	11.7%
30	Fort Bragg	Mendocino	Dec-12	7,273	11.7%
31	Alameda	Alameda	Jan-13	73,812	11.9%
32	Albany	Alameda	Jan-13	18,539	12.0%
33	Berkeley	Alameda	Jan-13	112,580	12.3%
34	Dublin	Alameda	Jan-13	46,036	12.4%
35	Emeryville	Alameda	Jan-13	10,080	12.4%
36	Fremont	Alameda	Jan-13	214,089	13.0%
37	Hayward	Alameda	Jan-13	144,186	13.4%
38	Livermore	Alameda	Jan-13	80,968	13.6%
39	Newark	Alameda	Jan-13	42,471	13.7%
40	Oakland	Alameda	Jan-13	390,724	14.8%
41	Piedmont	Alameda	Jan-13	10,667	14.8%
42	Pleasanton	Alameda	Jan-13	70,285	15.0%
43	San Leandro	Alameda	Jan-13	84,950	15.2%
44	Unincorporated Areas	Alameda	Jan-13	141,368	15.6%
45	Union City	Alameda	Jan-13	69,516	15.8%

	Jurisdiction	County	Effective Date	Population (2010)	Population w/Ban (%)
46	Ukiah	Mendocino	Jan-13	16,075	15.8%
47	Unincorporated Areas	Mendocino	Jan-13	59,081	16.0%
48	Laguna Beach	Orange	Jan-13	22,723	16.0%
49	Carmel	Monterey	Feb-13	3,722	16.0%
50	West Hollywood	Los Angeles	Feb-13	34,399	16.1%
51	Dana Point	Orange	Apr-13	33,351	16.2%
52	Belmont	San Mateo	Apr-13	25,835	16.3%
53	Brisbane	San Mateo	Apr-13	4,282	16.3%
54	Burlingame	San Mateo	Apr-13	28,806	16.4%
55	Colma	San Mateo	Apr-13	1,792	16.4%
56	Daly City	San Mateo	Apr-13	101,123	16.7%
57	Half Moon Bay	San Mateo	Apr-13	11,324	16.7%
58	Menlo Park	San Mateo	Apr-13	32,026	16.8%
59	Pacifica	San Mateo	Apr-13	37,234	16.9%
60	Portola Valley	San Mateo	Apr-13	4,353	16.9%
61	San Bruno	San Mateo	Apr-13	41,114	17.0%
62	South San Francisco	San Mateo	Apr-13	63,632	17.2%
63	Unincorporated Areas	San Mateo	Apr-13	88,362	17.4%
64	Capitola	Santa Cruz	Apr-13	9,918	17.4%
65	Santa Cruz	Santa Cruz	Apr-13	59,946	17.6%
66	Mountain View	Santa Clara	Apr-13	74,066	17.8%
67	San Mateo	San Mateo	Jun-13	97,207	18.0%
68	Glendale	Los Angeles	Jul-13	191,719	18.6%
69	San Carlos	San Mateo	Jul-13	28,406	18.6%
70	Los Altos	Santa Clara	Jul-13	28,976	18.7%
71	East Palo Alto	San Mateo	Oct-13	28,155	18.8%
72	Redwood City	San Mateo	Oct-13	76,815	19.0%
73	Cupertino	Santa Clara	Oct-13	58,302	19.2%
74	Huntington Beach	Orange	Nov-13	189,992	19.7%
75	Mill Valley	Marin	Nov-13	13,903	19.7%
76	Culver City	Los Angeles	Dec-13	38,883	19.8%
77	El Cerrito	Contra Costa	Jan-14	23,549	19.9%
78	Pittsburg	Contra Costa	Jan-14	63,264	20.0%
79	Richmond	Contra Costa	Jan-14	103,701	20.3%
80	San Pablo	Contra Costa	Jan-14	29,139	20.4%
81	Los Angeles	Los Angeles	Jan-14	3,792,621	30.6%
82	South Lake Tahoe	El Dorado	Jan-14	21,403	30.6%
83	Campbell	Santa Clara	Jan-14	39,349	30.7%
84	Arcata	Humboldt	Feb-14	17,231	30.8%
85	Los Gatos	Santa Clara	Feb-14	29,413	30.9%
86	Santa Barbara City	Santa Barbara	Mar-14	88,410	31.1%
87	Morgan Hill	Santa Clara	Apr-14	37,882	31.2%
88	Truckee	Nevada	Jun-14	16,180	31.2%
89	Beverly Hills	Los Angeles	Jul-14	34,109	31.3%
90	Davis	Yolo	Jul-14	65,622	31.5%
91	Walnut Creek	Contra Costa	Sep-14	64,173	31.7%
92	Tiburon	Marin	Sep-14	8,962	31.7%
93	Desert Hot Springs	Riverside	Sep-14	25,938	31.8%
94	Cloverdale	Sonoma	Sep-14	8,618	31.8%
95	Cotati	Sonoma	Sep-14	7,265	31.8%

	Jurisdiction	County	Effective Date	Population (2010)	Population w/Ban (%)
96	Healdsburg	Sonoma	Sep-14	11,254	31.9%
97	Petaluma	Sonoma	Sep-14	57,941	32.0%
98	Rohnert Park	Sonoma	Sep-14	40,971	32.1%
99	Santa Rosa	Sonoma	Sep-14	167,815	32.6%
100	Sebastopol	Sonoma	Sep-14	7,379	32.6%
101	Sonoma	Sonoma	Sep-14	10,648	32.6%
102	Unincorporated Areas	Sonoma	Sep-14	146,006	33.0%
103	Windsor	Sonoma	Sep-14	26,801	33.1%
104	San Rafael	Marin	Sep-14	57,713	33.2%
105	South Pasadena	Los Angeles	Oct-14	25,619	33.3%
106	Novato	Marin	Oct-14	51,904	33.4%
107	Sausalito	Marin	Oct-14	7,061	33.5%
108	Larkspur	Marin	Nov-14	11,926	33.5%
109	Indio	Riverside	Nov-14	76,036	33.7%
110	Palm Springs	Riverside	Nov-14	44,552	33.8%
111	Santa Clara	Santa Clara	Dec-14	116,468	34.1%
	Statewide			37,253,965	

Source: Author's calculations. Population statistics come from U.S. Census Bureau, Census 2010.

Table A.2: Event Study Regression Output for Figure 3a

		$Y_{sjw} = ln(TxnDuration)$			
D_{-24}	-0.004	D_{-5}	-0.006	D_{15}	0.018
	(0.011)		(0.013)		(0.015)
D_{-23}	-0.008	D_{-4}	0.008	D_{16}	0.019
	(0.014)		(0.013)		(0.014)
D_{-22}	0.002	D_{-3}	-0.005	D_{17}	0.026*
	(0.013)		(0.012)		(0.015)
D_{-21}	-0.016	D_{-2}	0.006	D_{18}	0.029*
	(0.015)		(0.011)		(0.016)
D_{-20}	0.003	D_{0}	0.033***	D_{19}	0.012
	(0.012)		(0.009)		(0.014)
D_{-19}	-0.008	D_{1}	0.025**	D_{20}	0.027*
	(0.016)		(0.011)		(0.015)
D_{-18}	0.001	D_{2}	0.045***	D_{21}	0.025
	(0.012)		(0.009)		(0.017)
D_{-17}	0.001	D_{3}	0.017	D_{22}	0.024*
	(0.013)		(0.014)		(0.012)
D_{-16}	-0.008	D_{4}	0.048***	D_{23}	0.024
	(0.016)		(0.012)		(0.016)
D_{-15}	0.000	D_{5}	0.028**	D_{24}	0.023
	(0.015)		(0.013)		(0.015)
D_{-14}	0.001	D_{6}	0.029**		
	(0.015)		(0.013)		
D_{-13}	-0.018	D_{7}	0.027*		
	(0.012)		(0.016)		
D_{-12}	-0.003	D_{8}	0.026*		
	(0.012)		(0.014)		
D_{-11}	0.002	D_{9}	0.034**		
	(0.014)		(0.013)		
D_{-10}	-0.002	D_{10}	0.018		
	(0.013)		(0.015)		
D_{-9}	-0.008	D_{11}	0.011		
	(0.015)		(0.015)		
D_{-8}	-0.004	D_{12}	0.008		
	(0.012)		(0.014)		
D_{-7}	0.003	D_{13}	0.023**		
	(0.014)		(0.009)		
D_{-6}	-0.004	D_{14}	0.024**		
	(0.011)		(0.010)		
		Num of Obs.		8673	
		Standard Errors		Cluster	
		Store FE		Yes	
		Week-of-sample FE		Yes	

Note: The table presents the results from event study Equation 1, as plotted in Figure 3a. The dependent variable is logged average transaction duration, measured in minutes, in store s, jurisdiction j, and week-of-sample w. Standard errors are in parentheses, estimated using two-way error clustering at the policy jurisdiction and week-of-sample level. Asterisks indicate: $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

Table A.3: Correlates of Paper Bag Choice (*Customer-Month Averages*)

	(1) Purchasing Paper (share)		(2) Purchasing Paper (share)	
	Coeff.	Std. Err.	Coeff.	Std. Err.
Customer-Month Variables				
Q2: 4-8 Items Scanned (share)	0.127***	0.006	0.114***	0.006
Q3: 9-18 Items Scanned (share)	0.176***	0.007	0.155***	0.007
Q4: > 18 Items Scanned (share)	0.138***	(0.005)	0.126***	(0.006)
Amount/Item ($)	0.004***	(0.000)	0.001***	(0.000)
Alcohol & Tobacco (share)	0.011***	(0.002)	0.002	(0.002)
Bakery & Deli (share)	0.004**	0.001	0.010***	0.001
Dairy & Refrigerated Items (share)	0.000	(0.003)	0.016***	(0.001)
Floral (share)	-0.013***	(0.004)	-0.009***	(0.002)
Prepared Foods (share)	0.037***	(0.002)	0.020***	(0.001)
Frozen Items (share)	-0.001	(0.003)	0.009***	(0.001)
Meat & Seafood (share)	0.024***	(0.002)	0.021***	(0.002)
Fresh Produce (share)	-0.011***	(0.002)	0.017***	(0.002)
Shelf-Stable Food Items	0.022***	(0.002)	0.014***	(0.001)
Baby Items (share)	0.041***	(0.005)	0.002	(0.004)
Pet Items (share)	-0.000	(0.003)	-0.012***	(0.002)
Full-service Lane (share)	-0.017	(0.010)	-0.010	(0.008)
Express Lane (share)	0.001	(0.009)	0.005	(0.008)
Trips/Month (#)	-0.002	(0.001)	-0.002***	(0.000)
Store Variables				
Median Income $55-72K (=1)	0.027**	(0.010)		
Median Income > $72K (=1)	0.051***	(0.014)		
Household Size (#)	-0.010	(0.020)		
White (share)	-0.256**	(0.112)		
Black (share)	-0.207*	(0.119)		
Asian (share)	-0.537***	(0.112)		
Over 65 (share)	-0.073	(0.110)		
Do not own vehicle (share)	0.416*	(0.230)		
Remodel Date (year)	-0.000	(0.001)		
Store Open Date (year)	-0.000	(0.000)		
Size (1000 ft^2)	-0.001***	(0.000)		
Self-checkout (=1)	0.008	(0.013)		
Urban (=1)	-0.005	(0.009)		
N	623,482		619,466	
R squared	0.039		0.378	
Standard Errors	Cluster		Cluster	
Customer FE	No		Yes	
Month-of-Sample FE	Yes		Yes	
Months-since-Policy FE	Yes		Yes	

Note: The dependent variable is the share of customer's i transactions where paper bags were purchased in month m and store s. Asterisks indicate the following: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses, estimated using two-way error clustering at the policy jurisdiction and month-of-sample level. *Source*: Customer-month variables come from author's calculations using the scanner data. Store variables were provided by the retailer, or come from Gicheva et al. (2010), who use 2000 US Census data for each store's census block-group. Urban areas are locations with populations densities greater than 500 people per square mile.

Table A.4: Effect of DCB Policies on Share of Transactions Purchasing Item Group (*Store-Week Averages*)

	(1) Produce	(2) Meat and Seafood	(3) Dairy and Refrigerated	(4) Frozen	(5) Bakery and Deli
Ban Effective Dummy	-0.003	-0.002	-0.002	-0.000	0.002
	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)
Mean Y_{sw} (2011)	0.61	0.34	0.47	0.26	0.29
	(6) Shelf-Stable Food	(7) Alcohol and Tobacco	(8) Baby Items	(9) Floral	(10) Pet
Ban Effective Dummy	-0.001	0.001	-0.000	-0.000	-0.000
	(0.002)	(0.001)	(0.000)	(0.001)	(0.000)
Mean Y_{sw} (2011)	0.73	0.19	0.02	0.04	0.05
Num of Obs.	8673	8673	8673	8673	8673
Standard Errors	Cluster	Cluster	Cluster	Cluster	Cluster
Store FE	Yes	Yes	Yes	Yes	Yes
Week-of-sample FE	Yes	Yes	Yes	Yes	Yes

Note: The table presents the results from difference-in-differences Equation 2. The dependent variable is 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. Standard errors are in parentheses, estimated using two-way error clustering at the policy jurisdiction and week-of-sample level. Asterisks indicate: $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

Table A.5: Average Transaction Level Characteristics in 2013 (*In-store Data*)

	(1)	(2)	(3)	(4)	(5)	(6)
	National Chain			Discount Chain		
	No DCB Policy	DCB Policy	Diff.	No DCB Policy	DCB Policy	Diff.
Bagger Present (=1)	0.74	0.69	0.048	0.47	0.08	0.392***
Checkout Interruption (=1)	0.06	0.07	-0.007	0.03	0.03	-0.007
Male (=1)	0.41	0.43	-0.021	0.42	0.53	-0.115***
White (=1)	0.65	0.71	-0.059	0.39	0.45	-0.054
Black (=1)	0.14	0.09	0.055*	0.28	0.31	-0.026
No Bag (=1)	0.05	0.15	-0.101***	0.06	0.24	-0.185***
Plastic Bag (=1)	0.86	0.00	0.865***	0.90	0.00	0.897***
Paper Bag (=1)	0.05	0.38	-0.331***	0.00	0.12	-0.115***
Bought Reus. Bag (=1)	0.00	0.02	-0.019**	0.00	0.19	-0.186***
Brought Reus. Bag (=1)	0.11	0.55	-0.440***	0.07	0.51	-0.446***
Plastic Bag (#)	4.03	0.00	4.030***	3.30	0.00	3.301***
Paper Bag (#)	0.16	0.75	-0.589***	0.00	0.28	-0.282***
Bought Reus. Bag (#)	0.00	0.03	-0.032**	0.00	0.31	-0.308***
Brought Reus. Bag (#)	0.26	1.22	-0.968***	0.09	0.88	-0.793***
Transaction Duration (min.)	1.49	1.72	-0.236**	1.31	1.59	-0.278***
N Obs.	333	157		631	156	

Note: Since 2013 is the pre-policy period for the treated stores, I group the treated and no-policy stores together (*No DCB policy*) and the prior-policy stores (*DCB policy*). Asterisks indicate the following: * p < 0.10, ** p <0.05, *** p <0.01. *Source*: Author's calculations from observational data collected in-store at checkout.

Table A.6: Effect of DCB Policies on Transaction Duration (*Scanner vs. In-store Data*)

	(1)	(2)	(3)	(4)
		National Chain		Discount Chain
	Scanner Data	Matched Data	In-store Data	In-store Data
DCB Effective (=1)	0.104***	0.099	0.155**	0.335***
	(0.026)	(0.0907)	(0.062)	(0.050)
Items Scanned (#)	0.024***	0.009***		
	(0.001)	(0.002)		
Amount Spent ($)	0.003***	0.002***		
	(0.002)	(0.001)		
Bagger Present (=1)		-0.062**	-0.092***	0.003
		(0.028)	(0.024)	(0.034)
Interruption (=1)		0.561***	0.576***	0.586***
		(0.079)	(0.066)	(0.052)
Male (=1)		-0.083***	-0.107***	-0.139***
		(0.021)	(0.021)	(0.019)
White (=1)		-0.060	-0.019	0.041
		(0.053)	(0.024)	(0.024)
Black (=1)		0.054	0.034	0.000
		(0.060)	(0.038)	(0.035)
No Bags (=1)		-0.197***	-0.055	-0.157***
		(0.056)	(0.048)	(0.034)
Plastic Bag (#)		0.035***	0.083***	0.105***
		(0.005)	(0.005)	(0.005)
Paper Bag (#)		0.088***	0.219***	0.237***
		(0.015)	(0.012)	(0.022)
Bought Reusable (#)		0.183***	0.354***	0.315***
		(0.059)	(0.047)	(0.021)
Brought Reusable (#)		0.096***	0.226***	0.243***
		(0.018)	(0.009)	(0.011)
Num of Obs.	22,061	687	1,692	2,228
Standard Errors	Cluster	Cluster	Cluster	Cluster
Store FE	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes
Register FE	Yes	Yes	Yes	Yes
Cashier FE	Yes	No	No	No
Hour FE	Yes	No	No	No
Mean Y_{tsd}	1.838	1.766	1.718	1.446

Note: The table presents the results from difference-in-differences Equation 11. The dependent variable is logged average transaction duration, measured in minutes, for transaction t in store s and date d. Standard errors are estimated using error clustering at the store-day level. Asterisks indicate the following: $*p < 0.10$, $**p < 0.05$, $***p < 0.01$. For the national chain, column (1) presents the model with the full sample of scanner data, column (3) presents the model with the full sample of in-store observational data, and column (2) presents the model with only the transactions matched between the scanner and in-store data. For the discount chain, column (4) presents the model with the full sample of in-store data.

Table A.7: Average Store and Transaction Characteristics from Pre-Policy Period (*DC Data*)

	Control	Treat	Diff.
Building Characteristics			
Building Size (ft^2)	42,990.17	45,927.67	-2937.500
Open Year	1985	1981	4.000
Register Characteristics			
Self-checkout (share)	0.17	0.33	-0.167
Full-service Registers (#)	4.83	6.00	-1.167
Express Registers (#)	4.08	3.67	0.417
Demographic Characteristics			
Median Income ($)	$67,379.92	$58,345.83	9034.083
Household Size (#)	2.41	2.20	0.203
White (share)	0.58	0.40	0.178
Black (share)	0.26	0.53	-0.275
Asian (share)	0.08	0.02	0.057**
Over 65 (share)	0.11	0.15	-0.046**
Do not own vehicle (share)	0.11	0.27	-0.166***
Urban (share)	1.00	1.00	0.000
Transaction Length (minutes)			
Full Sample	1.80	2.02	-0.224**
Smallest Txns. (scans < 8)	1.23	1.29	-0.062
Largest Txns. (scans $\geq$ 8)	2.42	2.73	-0.314**
Items Scanned (#)			
Full Sample	10.98	12.08	-1.093
Smallest Txns. (scans < 8)	3.72	3.70	0.020
Largest Txns. (scans $\geq$ 8)	18.98	20.18	-1.201
Amount Paid ($)			
Full Sample	32.77	35.04	-2.274
Smallest Txns. (scans < 8)	12.67	11.88	0.791*
Largest Txns. (scans $\geq$ 8)	54.65	57.33	-2.678
N Stores	12	6	–

Note: Asterisks indicate the following: * p < 0.10, ** p <0.05, *** p <0.01. *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. Transaction characteristics obtained from author's calculations using the scanner data.

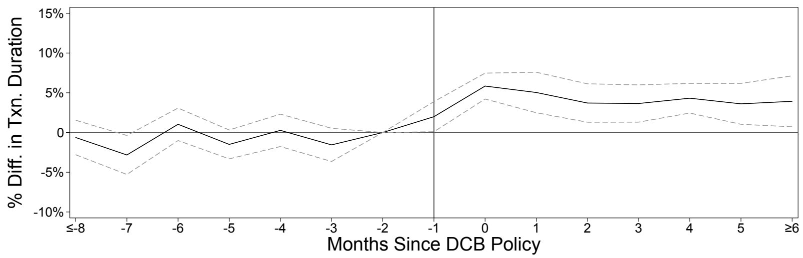
Table A.8: Store Heterogeneity: Effect of DCB Policies on Transactions per Shift (*Store-Week Averages*)

	Log Transactions per Shift (#)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ban Effective	-0.032***	-0.028***	-0.012	-0.035**	-0.030**	-0.017	-0.040***	0.001
	(0.010)	(0.010)	(0.013)	(0.014)	(0.011)	(0.016)	(0.013)	(0.025)
Ban Effective × Large Building Size		-0.006						-0.011
		(0.016)						(0.018)
Ban Effective × Median Income \$55–72K			-0.024					-0.041*
			(0.018)					(0.021)
Ban Effective × Median Income > \$72K			-0.041***					-0.073***
			(0.014)					(0.025)
Ban Effective × High Asian Share				0.006				0.042**
				(0.016)				(0.018)
Ban Effective × High Black Share					-0.004			-0.037**
					(0.017)			(0.016)
Ban Effective × Urban						-0.017		0.005
						(0.019)		(0.021)
Ban Effective × Low Vehicle Owner Share							0.017	-0.000
							(0.016)	(0.020)
Num of Obs.	8673	8673	8673	8673	8673	8673	8673	8673
R squared	0.959	0.959	0.960	0.959	0.959	0.959	0.959	0.960
Standard Errors	Cluster	Cluster	Cluster	Cluster	Cluster	Cluster	Cluster	Cluster
Covariates X_{sw}	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Store FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Week-of-Sample FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

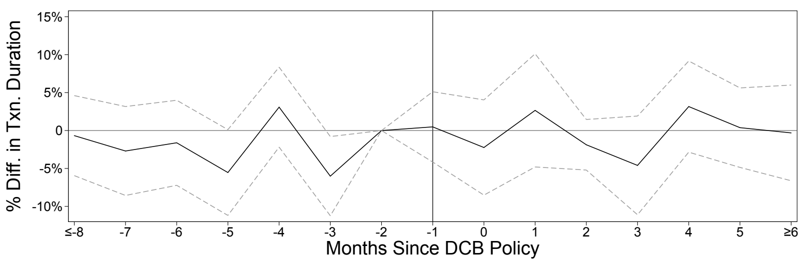
Note: The table presents the results from Equation 2. The outcome variable is the logged average number of transactions completed per 3 hour shift in store s, jurisdiction j, and week w. Column (1) replicates the log specification from column (1) of Table 4. Columns (2) through (7) include interactions of the ban effective dummy and dummies for store s being (i) above median with respect to building size, (ii) in a census block group with median income either $<$ \$55K (*omitted*), \$55–77K, or $>$ \$72K, (iii) in a census block group with an above median share of Asian residents, (iv) in a census block group with an above median share of Black residents, (v) in an urban census block group, and (vi) in a census block group with lower than median vehicle ownership. Column (8) includes all interactions with the ban effective dummy. Standard errors are estimated using two-way error clustering at the policy jurisdiction and week-of-sample level. Asterisks indicate the following: $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

Figure A.1: Effect of DCB Policies on Transaction Duration (*Customer-Month Averages*)

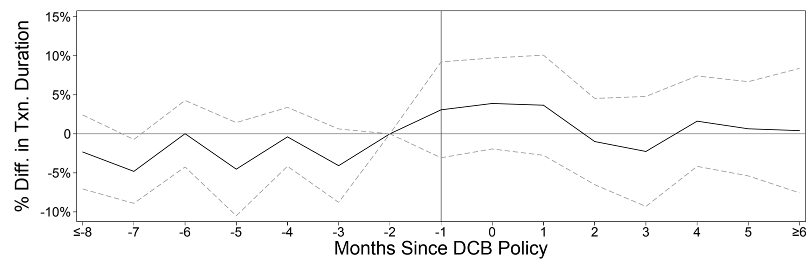
(a) All Households



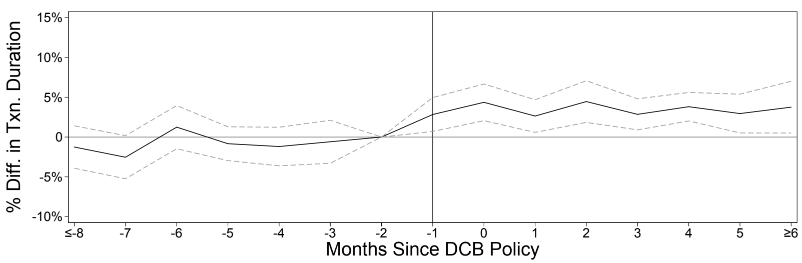
(b) Never Purchased Paper, Small Txn. Size. (Scans < 8)



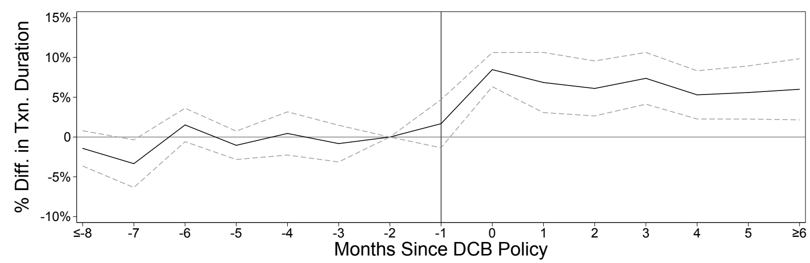
(c) Purchased Paper, Small Txn. Size. (Scans < 8)



(d) Never Purchased Paper, Large Txn. Size. (Scans ≥ 8)



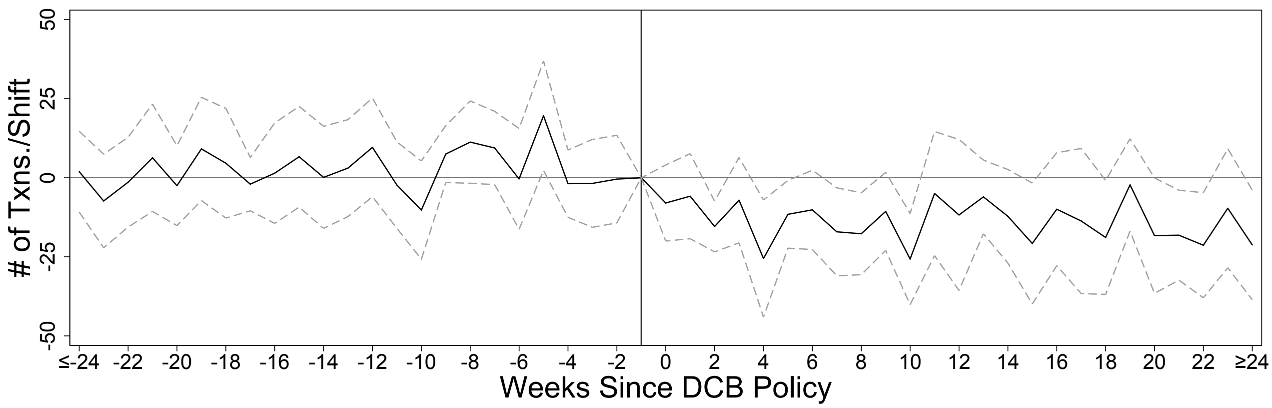
(e) Purchased Paper, Large Txn. Size. (Scans ≥ 8)



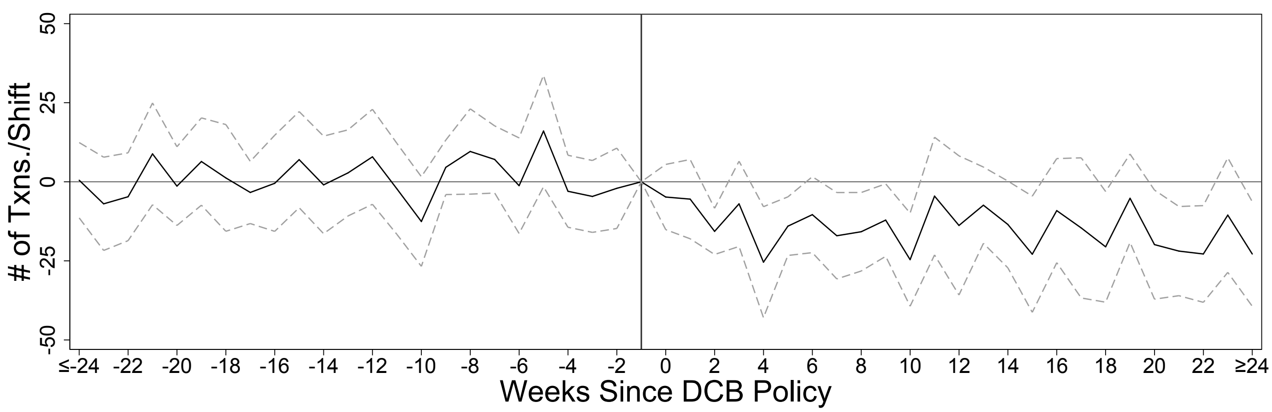
Note: Figure presents the full specification of event study Equation 9, with customer and month-of-sample fixed effects and control variables for the average number of items purchased per transaction, the amount spent per transaction, the types of items purchased, and months in sample—by customer i in store s, jurisdiction j, and month-of-sample m. The dependent variable is logged average transaction duration, measured in minutes, for customer i in store s, jurisdiction j, and month-of-sample m, for the entire sample of customers (panel a) and for all control customers compared to subsets of treated customers by transaction size and paper bag use (panels b–e). Upper and lower 90% confidence intervals are depicted in gray, estimated using two-way cluster robust standard errors on policy jurisdiction and week-of-sample.

Figure A.2: Effect of DCB Policies on Number of Transactions Completed per Shift (*Store-Week Averages*)

(a) Transactions per Shift—Without Control Variables



(b) Transactions per Shift—With Control Variables



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. Upper and lower 90% 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—alcohol and tobacco, floral department items, fresh meat and seafood, fresh produce, pet items, and baby items—in store s, jurisdiction j, and week-of-sample w.