Frictionless Grocery: How Cashierless Stores Reshape Shopping Trips *

Zhe Zhang[1], Hyoduk Shin[1], Simha Mummalaneni[2], Jonathan Z. Zhang[3], and Xiaofeng Liu[4]

[1] *University of California, San Diego*
[2] *William & Mary*
[3] *Colorado State University*
[4] *Baruch College*

September 14, 2026

Abstract

Cashierless stores are moving from experimentation to commercial deployment, enabled by advances in computer vision, sensor fusion, RFID, and AI-based checkout. This development is especially important in grocery retail: grocery is one of the largest and most frequent retail categories, yet traditional delivery-based e-commerce remains costly for small, perishable, and urgent purchases. Using individual-level transaction data from a major European convenience retailer, we study how consumers change their shopping behavior after adopting cashierless "CLS" stores. The data track purchases across the retailer's CLS and traditional stores, including basket composition, timing, prices, quantities, and product attributes. We estimate a difference-in-differences model that exploits variation in usage intensity among adopters with similar pre-adoption shopping behavior. CLS adoption increases shopping frequency, but spending does not rise proportionally. Instead, consumers fragment demand into smaller baskets, buy lower-priced items and smaller package sizes, and shift trips toward morning hours when time pressure is likely high. Spending in some restricted categories, particularly strong alcohol, also declines, suggesting that trip reallocation can disrupt some habitual shopping occasions. The findings show that cashierless stores are not merely faster checkout technology; they create a distinct grocery shopping occasion, with implications for cashierless-store deployment, assortment, replenishment, and channel strategy.

Keywords: Retail, Grocery, Cashierless Technology, Just Walk Out, Automation

*We thank Vladimir Atanasov for helpful comments and feedback. Email addresses: z9zhang@ucsd.edu, hdshin@ucsd.edu, nmummalaneni@wm.edu, jonathan.zhang@colostate.edu, and xiaofeng.liu@baruch.cuny.edu

1 Introduction

Grocery is one of the largest, most universal, and least digitized domains of retail. In the United States alone, supermarkets generate approximately $1 trillion in sales, and food accounts for 12.9% of household expenditures, ranking behind only housing and transportation (FMI, 2026; Sinclair and Stewart, 2026). Unlike many other industries, grocery has not been transformed by e-commerce. Purchases are frequent, local, perishable, and often immediate: coffee before work, a drink on the way to a meeting, a snack between errands, or a missing ingredient for dinner. These occasions are economically difficult for delivery-based e-commerce to serve because the retailer must absorb picking, substitution management, cold-chain handling, delivery windows, and last-mile delivery costs; indeed, picking is typically the largest aggregate cost in online grocery fulfillment (McKinsey & Company, 2026). As a result, physical stores remain central to grocery retail even as digital grocery channels grow.

Cashierless stores offer a different path. Rather than replacing the store visit with delivery, they digitize and automate the physical store itself. Consumers scan in, pick up products, and leave. Purchases are detected automatically through computer vision, sensors, RFID, and AI-based checkout systems, and payment occurs in the background (Amazon Web Services, 2026). The format preserves the immediacy of physical retail while removing one of its most visible frictions: checkout. This makes cashierless grocery neither traditional brick-and-mortar retail nor conventional e-commerce, but a hybrid format potentially suited to small, fast, immediate grocery occasions.

Yet the consequences of introducing a cashierless format are difficult to predict ex ante. Prior research on shopping friction suggests that reducing waiting and transaction costs should make shopping easier and may generate incremental demand by enabling trips that consumers would otherwise forgo (Allon et al., 2011; Campbell and Frei, 2010; Xue et al., 2011). At the same time, omnichannel and store-entry research shows that a new format can also redirect purchases that would otherwise have occurred elsewhere in the retailer's network, creating cannibalization rather than demand expansion (Avery et al., 2012; Shriver and Bollinger, 2022). Even where prior theory suggests a directional effect, it provides limited guidance about magnitude or about whether more usage translates into more spending. Cashierless stores could therefore expand demand, reallocate existing demand, or produce both effects simultaneously.

The uncertainty extends beyond aggregate revenue. By lowering the fixed cost of a store visit, cashierless retail may change how consumers organize their shopping. Consumers may shop more frequently and spend more overall, or they may simply fragment relatively fixed demand across more frequent, smaller baskets. The format may also change the composition of those baskets. If consumers shop closer to the moment of need, they may shift toward smaller package sizes, ready-to-consume products, or categories associated with quick convenience missions rather than planned or stock-up purchases. Over time, such changes could reshape established shopping routines. In short, the format may change the shopping occasion itself. These responses may vary across consumers depending on their existing shopping frequency, routines, and time constraints. Thus, neither the average effect nor the underlying behavioral changes are obvious a priori.

These ambiguities raise four central questions. First, does cashierless retail increase the retailer's overall revenue and purchasing activity? Second, to what extent does this activity substitute for shopping at traditional stores? Third, how does adoption change trip frequency, timing, basket size, and composition? Fourth, how do these effects differ across consumer segments? Scant existing research leaves limited field evidence on these outcomes that are of interest to firms and business scholars.

We study these questions using individual-level transaction data from a major European convenience retailer that introduced cashierless "CLS" stores. The data track consumers' purchases across both CLS and traditional stores, including basket composition, purchase timing, prices, quantities, store format, and product attributes. Because the same consumers are observed at both CLS and traditional stores, we can measure how adoption changes shopping across the retailer's broader store network, and we can compare the same consumer's purchases across the two formats. Settings where only the new format is observed do not allow these comparisons.

Our empirical design compares consumers who became regular CLS users with consumers who tried CLS once but did not incorporate the format into their regular shopping routines. This comparison is useful because both groups demonstrated awareness of, access to, and willingness to try the technology. We further match consumers on pre-adoption shopping behavior and estimate difference-in-differences models with individual and week fixed effects. Our primary specification uses continuous treatment intensity, measured as each consumer's average number of CLS trips per post-adoption week. Because consumers choose whether and how intensively to use CLS stores, our primary object of interest is an average treatment effect on the treated (ATT). Specifically, our goal is to quantify how post-adoption changes in shopping behavior scale with sustained CLS usage among adopters. In the Web Appendix, we also use a shift-share instrumental variable based on geographic access to CLS stores to provide complementary evidence.

Our results show that CLS stores do not simply make existing grocery trips faster. Instead, they create a distinct shopping occasion. Consumers who use CLS stores shop more frequently, and the time between shopping trips shortens. However, spending does not rise proportionally: each additional weekly CLS trip is associated with about 27% more weekly shopping trips but only 16% more weekly spending. Instead, consumers fragment demand into smaller baskets: they buy fewer items per trip, spend less per trip, purchase lower-priced items, and choose smaller package sizes. This pattern suggests that the format shifts consumers from larger, consolidated trips toward smaller, more immediate purchases.

The timing of shopping changes as well. The increase in trip frequency is concentrated in morning hours, especially 6:00–10:00am, when time pressure is likely to be high. This concentration is consistent with a time-pressure mechanism. The value of a fast, predictable shopping trip may be greatest when consumers face tight schedule constraints and checkout friction makes a conventional trip most costly in time.

The product-choice evidence supports the same interpretation. CLS adopters shift toward smaller-package and lower-volume purchases. This pattern is not explained by CLS stores carrying different assortments. In a within-consumer, within-product comparison, the same consumer buying the same base product chooses a smaller package at CLS stores than at traditional stores. The consumer is not globally becoming a smaller-size buyer; rather, the CLS format induces a smaller, more immediate shopping occasion.

CLS adoption also changes behavior outside the CLS channel. Higher CLS usage is associated with fewer traditional-store trips and slightly smaller traditional-store baskets, suggesting substitution from existing store routines.

One finding is not predicted by either demand expansion or simple channel substitution. While CLS stores sell no tobacco or alcohol, consumers retain access to the same traditional stores they previously visited. It may be theorized that a new store option should leave these categories roughly unchanged. Instead, spending in them falls substantially, by approximately 15% per weekly CLS trip. This is most clear in strong alcohols (spirits), which is associated with a decline of roughly a third among spirits buyers for each weekly CLS trip. We interpret this result cautiously, but it suggests that moving trips into a cashierless convenience format can disrupt habitual shopping occasions tied to traditional-store routines.

These findings help clarify the strategic role of cashierless retail. The industry trajectory has been mixed: Amazon has scaled back its cashierless Amazon Go grocery stores, but they have shifted to licensing their underlying Just Walk Out technology to third-party locations such as airports, stadiums, hospitals, and small-format stores (Kumar, 2024; Bensinger, 2024). Operators of small-format cashierless stores similarly emphasize speed, immediate access, and 24/7 availability, with industry reports citing average visits of 60 seconds or less and customer satisfaction levels substantially above typical stores (Briggs, 2023). Our results suggest why the format may be more compelling in curated, mission-driven, small-basket environments than in large stock-up grocery settings. Cashierless technology appears most valuable as an enabler of small, time-sensitive grocery occasions rather than as a universal replacement for supermarkets.

This paper makes three contributions. First, we contribute to research on AI-enabled and self-service retail technologies by providing field evidence on realized purchase behavior rather than adoption intentions or attitudes. Prior work on autonomous and self-service retail has studied adoption, satisfaction, convenience, privacy, trust, and service tradeoffs (Benoit et al., 2024; Bulmer et al., 2018; Cui et al., 2022; Hazée et al., 2025; Meuter et al., 2000; Nusrat and Huang, 2024; Pillai et al., 2020). We show that beyond the service interface, cashierless technology changes shopping frequency, timing, basket size, and product choice.

Second, we contribute to omnichannel and store-format research by studying a hybrid channel: a physical store with a digital and AI-enabled operating model. Prior omnichannel research often asks whether new channels complement or cannibalize existing channels (Avery et al., 2012; Bell et al., 2015, 2018; Fisher et al., 2019; Neslin, 2022; Shriver and Bollinger, 2022; Wang and Goldfarb, 2017). We show that this framing is incomplete for cashierless convenience formats. Beyond adding or subtracting demand, CLS stores change what people buy, how much they buy, and when they buy. The relevant outcome is how consumers reorganize purchasing when they can buy in smaller, more immediate, and more time-constrained moments.

Third, we contribute to research on shopping friction and consumer time costs. Prior work shows that consumers are sensitive to waiting time, queueing, transaction costs, and shopping frictions (Allon et al., 2011; Belavina et al., 2017; Campbell and Frei, 2010; Ülkü et al., 2020; Xue et al., 2011). We show that eliminating checkout friction changes the structure of grocery demand. Consumers move from larger, more consolidated trips toward smaller, more frequent, just-in-time purchases. The morning concentration and package-size results provide behavioral evidence consistent with the view that the value of friction reduction is highest when time constraints are binding and when consumers seek immediate consumption rather than stock-up purchasing. In doing so, we connect the literature on shopping costs, retail formats, and purchase missions to a new AI-enabled physical grocery context (Aydinli et al., 2021; Benoit et al., 2019; Chintala et al., 2024; Jindal et al., 2020; Zhang et al., 2022).

The rest of the paper proceeds as follows. Section 2 reviews related research and develops the conceptual positioning. Section 3 describes the institutional setting and data. Section 4 presents the empirical strategy. Section 5 reports the main results on shopping frequency, spending, basket size, category composition, and time of day. Section 6 examines mechanisms related to time pressure, assortment, pricing, and package-size choice. Section 7 presents heterogeneity. Section 8 concludes with theoretical and managerial implications. Robustness checks and supporting analyses appear in the Web Appendix.

2 Related Literature

Our study connects three streams of research: self-service and autonomous retail technology, omnichannel and store-format strategy, and shopping friction and grocery shopping. The central gap across these literatures is that we know much more about technology adoption and channel availability than about how a fully autonomous physical grocery format changes actual shopping routines.

2.1 Self-service technology and autonomous retail

A large literature studies self-service technologies, including ATMs, online banking, self-checkout kiosks, mobile ordering, scan-and-go systems, and other technology-mediated service encounters. This research has shown that self-service technologies can improve perceived convenience, reduce waiting time, increase customer control, and alter the economics of service delivery. At the same time, they may reduce interpersonal interaction, shift effort from employees to customers, and affect loyalty, satisfaction, and trust (Bulmer et al., 2018; Campbell and Frei, 2010; Meuter et al., 2000; Nusrat and Huang, 2024; Xue et al., 2011). Much of this literature therefore frames self-service technology as a tradeoff between efficiency and service experience.

Cashierless stores represent a qualitatively different step in this evolution. Traditional self-checkout still requires consumers to scan products, resolve errors, and complete payment. Mobile ordering or self-service kiosks change the ordering interface but preserve a service encounter. Cashierless stores instead remove checkout as an explicit consumer task. The consumer does not choose a faster lane or interact with a kiosk; the consumer simply leaves. This distinction matters because the behavioral effect may extend beyond satisfaction with checkout. If the fixed time cost of entering a store falls enough, consumers may create new trips, shop at different times, and select different baskets.

Existing research on autonomous retail has largely focused on consumer perceptions, technology acceptance, convenience dimensions, privacy concerns, or service-design features. De Bellis and Johar (2020) identify barriers to consumer adoption of autonomous shopping systems. Pillai et al. (2020) study shopping intention at AI-powered automated retail stores. Cui et al. (2022) examine consumer response to AI-enabled checkout technologies. Benoit et al. (2024) show that consumer responses to autonomous stores depend on design features such as access, staff support, checkout, verification, convenience, autonomy, and safety. Hazée et al. (2025) examine reasons for and against consumer adoption of just-walk-out stores, including access convenience, transaction convenience, search convenience, no need for interaction, technology reliability concerns, privacy, and assortment concerns. These studies are important because autonomous formats require consumers to trust invisible systems that identify products and charge accounts correctly. However, adoption intentions cannot reveal whether consumers actually spend more, shift trips across formats, or change basket composition after adoption.

Our study complements this work by comparing observed purchase behavior before and after consumers begin using cashierless stores. This distinction is important for evaluating the promise of AI-enabled retail. Industry narratives often emphasize checkout speed, labor savings, and customer satisfaction (Kumar, 2024; Amazon Web Services, 2026; Briggs, 2023). Those are meaningful outcomes, but they do not directly answer the strategic question of how cashierless stores affect demand. We show that the technology's impact extends beyond faster checkout: it changes the shopping occasion itself.

2.2 Omnichannel retail, store entry, and format reallocation

Our paper also contributes to research on omnichannel retailing, store entry, and store-format strategy. A central question in this literature is whether adding a channel expands demand or cannibalizes existing channels. Prior work shows that channel additions can generate complementary effects when they improve information, access, trust, convenience, or fulfillment. Avery et al. (2012) show that adding physical stores to a direct retailer can create complex cross-channel effects over time. Bell et al. (2015, 2018) show that showrooms can increase demand and generate operational benefits for online retailers. Wang and Goldfarb (2017) show that offline stores can drive online sales. Fisher et al. (2019) study the value of rapid delivery in omnichannel retailing. Neslin (2022) conceptualizes omnichannel retailing as a continuum integrating online and offline channels along the customer journey.

At the same time, channel additions can substitute for existing demand when the new channel overlaps with current formats. Shriver and Bollinger (2022) show that retail store entry can generate both demand expansion and cannibalization, depending on travel costs and consumer demand patterns. Store-entry research similarly shows that new store openings can substantially change consumer store choice and incumbent performance. Singh et al. (2006), for example, study Walmart Supercenter entry and show that entry affects incumbent supermarket sales primarily through changes in visit frequency. Zhang et al. (2022) show that physical stores create value in part through product inspection, suggesting that channel value depends on the consumer task and product context. Chintala et al. (2024) show that online and offline grocery baskets differ systematically, highlighting how channel environments shape basket composition.

Cashierless CLS stores differ from many of the channel additions studied in prior work. They are physical stores with a digital operating model. They have the immediacy of brick-and-mortar retail, but the frictionless data infrastructure of digital retail. They also differ from large-format stores because they are optimized for speed, accessibility, and immediate consumption rather than assortment breadth or stock-up purchasing.

This hybrid nature makes the demand consequences theoretically ambiguous. CLS stores may expand demand by serving moments that traditional stores cannot profitably serve, cannibalize small purchases that would otherwise occur at traditional stores, or, as our findings support, reallocate demand across shopping occasions. The key change is in how the consumer organizes purchasing over time.

This perspective extends omnichannel and store-entry research by moving from channel substitution to occasion reconfiguration. In conventional omnichannel settings, the consumer chooses among channels to fulfill a purchase goal. In our setting, the new format may change the purchase goal itself: a planned purchase becomes a quick top-up, or a larger bottle becomes a single-serve drink. Thus, channel strategy and store-format strategy cannot be evaluated only through incremental revenue. They must also consider how a new format reorganizes trips, baskets, and routines.

2.3 Shopping friction, time pressure, and grocery shopping occasions

The third relevant stream of research examines how shopping costs shape consumer behavior. Consumers face fixed and variable costs when shopping: travel time, search time, waiting time, payment time, mental effort, and the risk of delays. When these costs are high, consumers may consolidate purchases into larger baskets, delay consumption, or avoid trips altogether. When these costs fall, consumers may shop more frequently and closer to the moment of need.

Prior research shows that consumers are sensitive to waiting and transaction frictions. Allon et al. (2011) estimate the value of reductions in wait time in fast-food drive-through settings. Ülkü et al. (2020) show that queueing and waiting experiences affect consumption. Belavina et al. (2017) show that online grocery subscription models can lead to smaller and more frequent orders. Campbell and Frei (2010) and Xue et al. (2011) show that self-service banking channels can increase transaction frequency while changing customer profitability and retention. These findings suggest that lowering transaction costs can alter consumer behavior even when the product need remains similar.

Grocery provides an especially important setting for these mechanisms. Grocery shopping consists of multiple missions, ranging from large stock-up trips to smaller fill-in and immediate-consumption trips. Traditional online grocery tends to work best when fixed ordering, picking, and delivery costs can be spread over larger planned baskets. By contrast, many grocery needs are small and immediate. These needs are difficult for delivery-based e-commerce to fulfill profitably, but they are well suited to a small physical format that is fast, local, and highly accessible.

Cashierless CLS stores provide a strong setting for studying these mechanisms because the technology eliminates one salient component of shopping friction: checkout. The fixed time cost of a trip matters most when the consumer's opportunity cost of time is high. A two-minute wait may be trivial during an unhurried evening trip but costly during a morning commute. This logic implies that the effect of cashierless retail should not be uniform across the day. It should be strongest when time pressure is most binding. Our time-of-day results are consistent with this view. CLS usage shifts trips toward morning hours, when consumers are likely to face schedule constraints.

This logic also predicts smaller baskets and smaller packages. If consumers can shop closer to the moment of need, they need not stockpile. They may buy one drink rather than a multi-pack, a small snack rather than a larger grocery basket, or a ready-to-eat item rather than a planned pantry item. Grocery shopping has long been characterized by different shopping missions, including stock-up and fill-in trips, and recent work shows that channel environments can systematically shape basket structure (Benoit et al., 2019; Chintala et al., 2024; Jindal et al., 2020). Our same-consumer, same-product size-choice result supports this occasion-specific interpretation. Consumers choose smaller sizes at CLS stores while continuing to choose larger sizes for the same products at traditional stores.

This logic also connects to research on impulse and vice purchasing. Easier access and reduced social friction may increase some immediate-consumption purchases, but changing the shopping format may also reduce exposure to categories that depend on habitual routines, display cues, or interpersonal scripts (Huyghe et al., 2017; Iyer et al., 2020; Massara et al., 2014; Sun et al., 2023). The restricted-category results in our setting show both sides of this tension. The same format that enables more frequent convenience trips can also disrupt purchases of categories like tobacco and alcohol that are absent from the CLS assortment.

Taken together, these literatures suggest that cashierless stores can act as a friction-reducing grocery format that reshapes the temporal structure of consumption. Our empirical contribution is to document this process using individual-level field data.

3 Data and Institutional Setting

3.1 Institutional Context

We study the introduction of cashierless convenience stores by the largest convenience store chain in Country A, operating over 9,000 locations as of our study period. In 2021, the chain began deploying "CLS" stores. These are small, unmanned retail units, most roughly the size of a shipping container, equipped with computer vision and sensor-based checkout technology. Customers enter by scanning the chain's already-popular loyalty app, select items from the shelves, and exit. The system automatically detects the items taken and charges the customer's account. No cashier or checkout station is involved.

During our study period (June 2021 to March 2022), the chain launched 24 CLS stores. While these locations are spread across eight cities, CLS activity is heavily concentrated in three major metropolitan areas—City A, City B, and City C—which together account for roughly 96% of all CLS transactions. The stores opened in distinct waves: an initial pilot of eight stores in June 2021, followed by batches of three in September 2021, four in December 2021, one in February 2022, and eight in March 2022. Three of the original pilot stores—all small, low-traffic sites—closed within their first few months and do not reappear in the data, while the remaining 21 stores stay open through the end of our sample.[1] Table 1 documents the resulting growth in CLS transaction volume and user base. The staggered rollout of stores and consumer adoption generates useful variation in both the timing and intensity of treatment.

Three features of this setting are important for our analysis. First, while CLS stores carry many of the same product categories as the chain's traditional stores—including beverages, snacks, and ready-to-eat meals—they offer a smaller overall assortment. Tobacco and alcohol are not sold at CLS stores, and some other categories (e.g., fresh bread) have limited availability due to the unstaffed format. The core convenience-store assortment for non-restricted categories is nonetheless preserved. Second, because the loyalty app is *required* for entry into CLS stores, all CLS transactions are captured in our data. Customers who never use CLS stores also appear in our data through

[1] Four of the early City A stores record no transactions in October–November 2021 before resuming in December, consistent with a brief temporary closure or a gap in data coverage during these two months. Because our identification relies on consumer-level adoption timing and treatment intensity rather than store-level calendar availability, and because these stores account for a small share of CLS activity in this window, this episode does not materially affect our estimates.

Month	CLS trips	Unique consumers
June 2021	24	21
July 2021	116	74
August 2021	96	64
September 2021	371	282
October 2021	2,485	1,453
November 2021	1,702	1,027
December 2021	2,448	1,307
January 2022	2,479	1,296
February 2022	2,710	1,306
March 2022	3,222	1,179

Table 1: CLS Store Rollout. Monthly CLS transaction counts and unique consumers using CLS stores in our estimation sample.

voluntary app usage at traditional stores, providing a comprehensive view of shopping behavior across formats. We do not regard the app requirement as a meaningful source of selection for our analysis. The consumers we study were already substantial app users at traditional stores before their first CLS trip, so we observe well-measured pre-CLS behavior rather than people who adopted the app specifically to use a CLS store.[2] Third, the CLS stores in our study were placed in new locations that the chain did not previously serve, and the existing traditional store network remained entirely unchanged during the rollout period. We discuss the implications of this bundled treatment—new location and new technology—for interpretation of our results in Section 8.

3.2 Transaction Data

Our primary data consist of individual item-level transactions recorded through the retailer's loyalty app. For each transaction, we observe the consumer identifier, date and time of purchase, store identifier and format (CLS vs. traditional), product description and category, unit price, quantity, and total spending. We define a shopping trip as a unique combination of consumer, store, and timestamp; i.e., all items purchased by the same consumer at the same store at the same time constitute one trip. This trip-level definition is our primary unit of analysis for shopping frequency.

We aggregate trips to the individual-week level, where weeks begin on Mondays. For each consumer-week, we compute the number of trips, total spending in local currency units (LCU),

[2] We account for this issue by implementing an activity filter for consumers in our panel. See Section 3.3 for more details about this restriction.

and total items purchased. To construct the analysis panel, we include all weeks between each consumer's first and last observed transaction, filling in zeros for weeks with no activity. This ensures that weeks of inactivity are properly represented, which is important for estimating effects on both the extensive margin (whether a consumer shops at all) and the intensive margin (how much they shop conditional on shopping). We restrict the panel to trips made between 6am and 11pm ("main hours"). A few CLS locations operate around the clock, but trips outside this window account for only 0.7% of the total. Table 2 summarizes the structure of the resulting panel dataset.

Statistic	Value
Total individuals	4,438
Treated consumers	2,219
Control consumers	2,219
Date range	2021-06-01 to 2022-03-27
Total calendar weeks	43
Individual-week observations (weekly panel)	170,249
Individual-week-period observations (time-of-day panel)	680,996
Pct zero-trip individual-weeks	27.3%
Total shopping trips (main hours, 6am–11pm)	363,797
Mean trips per consumer-week (unconditional)	2.14
Mean trips per consumer-week (conditional on at least one trip)	2.94

Table 2: Panel Structure and Sample Overview.

3.3 Sample Construction and Group Definitions

Our analysis focuses on consumers who adopted CLS stores during the study period; i.e., those who made at least one transaction at a CLS store. We define the *adoption date/week* for each consumer as the date/week of their first CLS trip. This event-based definition allows us to construct consumer-specific pre- and post-treatment periods, exploiting the staggered nature of adoption across individuals.

To ensure reliable estimation of pre- and post-adoption behavior, we require that each consumer have at least three weeks with shopping activity in both the 12-week window before and the 12-week window after their adoption date. This filter removes consumers who appear only sporadically in the data and for whom pre-treatment trends cannot be credibly estimated. The resulting sample retains consumers with sufficient longitudinal coverage on both sides of the adoption event.

Among the consumers passing the activity filter, we classify individuals into two groups based on their total CLS store usage during the study period:

Treated group: Consumers who made more than one CLS store trip in total. These consumers returned to CLS stores after their initial visit and incorporated the format into their shopping behavior to varying degrees.

Control group: Consumers whose entire CLS store history consists of a single trip. These individuals tried the technology—they scanned in, completed a purchase, and left—but never returned during the study period. We refer to this set of consumers as the "one-trip control" group.

The one-trip control group is a natural comparison set. These consumers demonstrated the same initial willingness to try a cashierless store as the treated group: they were aware of CLS stores, downloaded or already possessed the required app, and completed a transaction. Their decision not to return may reflect geographic inconvenience, commute patterns that do not pass a CLS location, or simply a lack of regular need, rather than a fundamental difference in openness to the technology. The within-adopter comparison thus mitigates selection concerns that would arise from comparing CLS adopters to consumers who never tried the format. We discuss this further in Section 4.

We further refine comparability using coarsened exact matching (CEM) on pre-treatment shopping behavior. Specifically, we match treated and control consumers on pre-adoption means of weekly trips, weekly spending, weekly items purchased, number of distinct stores visited, and the consumer's primary city. The matching is one-to-one within coarsened strata, ensuring that each treated consumer is paired with a control consumer who exhibited similar baseline shopping patterns in the same metropolitan area. In Section A of the Web Appendix, we provide a balance table which shows that the treatment and control groups are similar on observable criteria during the pre-treatment period. The one exception is a small difference in pre-treatment window length, which arises because the matching procedure does not match on adoption timing.

3.3.1 Early-adopter (pilot) exclusion

CLS's rollout began with a low-volume soft-launch before the format reached scale. We date the "real launch" to the first week in which system-wide CLS volume reached at least 50 trips (the week of 13 September 2021); earlier CLS transactions come from a handful of pilot stores operating at very low volume. We exclude *pilot adopters*—the 126 treated consumers whose first CLS trip predates the real launch—together with their CEM-matched controls (126 pairs in all), leaving 2,219 treated and 2,219 control consumers. Three considerations motivate this. First, pilot adopters have little or no clean pre-treatment window (a median of 6.9 pre-adoption weeks versus 24.9 for the retained sample; about one-fifth have fewer than four). Second, the pilot-era CLS experience may differ from the steady-state format the rest of the sample encountered (e.g., limited hours or assortment). Third, the pilot stores were largely transient. Of the eight CLS stores these adopters used during the pilot window, three (all low-volume) stopped recording transactions months before the panel ends, whereas none of the sixteen post-launch CLS stores did. Pilot adopters are few in number (5.4% of the treated group), so their inclusion has minimal effect on the estimates.

3.3.2 Treatment intensity

Rather than relying solely on the binary treated/control distinction, our primary specifications exploit continuous variation in treatment intensity. We define

$$\text{CLSUsage}_i = \frac{\text{Total CLS trips by consumer } i}{\text{Weeks observed after adoption week}}$$

which measures the average number of CLS trips per week in the post-adoption period. For control consumers (who made exactly one CLS trip), this measure is close to zero. For treated consumers, it ranges from modest usage (e.g., one CLS trip every few weeks) to frequent usage (multiple CLS trips per week).

Figure 1 shows the distribution of adoption timing across the study period for treated consumers. Adoption is spread across several months, reflecting the staggered CLS store openings and organic consumer discovery. This variation is important. Treated consumers enter sustained CLS usage at different calendar weeks, so treatment timing is separated from common-week shocks. The one-trip control group consists of consumers who never transition to sustained CLS usage, so they provide a low-exposure counterfactual throughout the panel. A large one-trip control group reduces reliance on already-treated consumers as implicit controls, the source of bias in staggered difference-in-differences designs (Goodman-Bacon, 2021). Section 4 discusses how we address this issue.

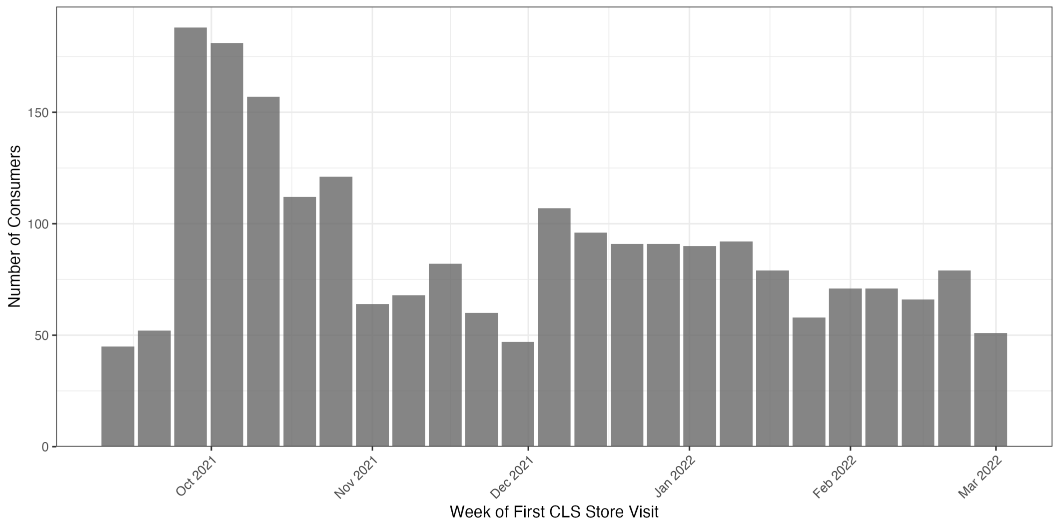

Figure 1: Distribution of Adoption Timing. The histogram shows the number of treated consumers by the week of their first CLS trip.

Variable	Definition
Y_{it}: Trips per week	Count of shopping trips by consumer i in week t
Y_{it}: Spending per week	Total gross spending (LCU) by consumer i in week t
Y_{it}: Items per week	Total items purchased by consumer i in week t
Post_{it}	$= 1$ if week t is after consumer i's first CLS-trip week
Treated_i	$= 1$ if consumer i made > 1 CLS trip (treated group)
CLSUsage_i	Total CLS trips $\div$ weeks observed post-adoption
Time period p	Morning (6–10), Work (10–16), Evening (16–20), Night (20–23)

Table 3: Key Variable Definitions.

3.3.3 Summary Statistics

Table 3 summarizes the key variables used throughout the analysis. Table 4 reports summary statistics for the key outcome variables at the individual-week level, disaggregated by group (treated vs. control) and period (pre- vs. post-adoption). Both groups exhibit substantial week-to-week variation in shopping activity, with a non-trivial fraction of zero-trip weeks reflecting the intermittent nature of convenience-store shopping. The post period overlaps substantially with winter months, as shown by the distribution of adoption timing in Figure 1. Thus, both groups show pre- vs. post-adoption differences in trips and spending. For the control group, these raw differences reflect common seasonal or calendar shocks rather than treatment. Week fixed effects absorb these common time-varying factors.

Panel	Trips/week	Spending/week (LCU)	Items/week	Pr(any trip)	N (indiv-weeks)
Treated, Pre-adoption	2.15 (2.40)	30.17 (44.74)	6.44 (8.94)	0.72	48,460
Treated, Post-adoption	2.15 (2.40)	29.86 (43.38)	6.09 (8.38)	0.74	34,562
Control, Pre-adoption	2.17 (2.41)	30.52 (43.05)	6.53 (8.69)	0.73	50,538
Control, Post-adoption	1.91 (2.25)	28.70 (44.87)	5.71 (8.36)	0.70	32,251

Table 4: Summary Statistics by Group and Period. Mean values of weekly shopping outcomes (standard deviations in parentheses), excluding the adoption week (relative week 0). Pr(any trip) is the fraction of individual-weeks with at least one trip.

4 Empirical Strategy

4.1 Empirical Model

Using the data and sample described in Section 3, we employ a difference-in-differences framework with continuous treatment intensity to estimate the effect of CLS store adoption on shopping behavior. Our baseline specification is:

$$Y_{it} = \alpha_i + \gamma_t + \beta_0 \cdot \text{Post}_{it} + \beta_1 \cdot \text{Post}_{it} \times \text{Treated}_i + \beta_2 \cdot \text{Post}_{it} \times \text{Treated}_i \times \text{CLSUsage}_i + \varepsilon_{it} \quad (1)$$

where Y_{it} represents outcome variables for individual i in week t, such as number of trips, total spending, or items purchased. The terms α_i and γ_t represent individual and week fixed effects, respectively. Post_{it} is an indicator equal to one for weeks after consumer i's first CLS trip. Treated_i

indicates whether consumer i is in the treatment group (consumers with more than one CLS trip). CLSUsage_i measures treatment intensity as the consumer's average number of CLS store trips per week in the post-adoption period.

The coefficient of primary interest is β_2, which measures how outcomes change with each additional CLS trip per week of sustained usage among adopters. Rather than estimating a single average effect for all treated consumers, this specification uses the continuous variation in CLS store usage described in Section 3. The coefficient β_2 represents the marginal effect of one additional CLS store trip per week on the outcome, comparing consumers with different sustained usage levels but similar pre-treatment behavior. Treatment intensity varies substantially among treated consumers—some visit CLS stores multiple times per week, while others use them only occasionally—so a single binary treated/control contrast would average over economically different treatments. Because consumers choose whether and how intensively to use CLS stores, our primary object of interest is an average treatment effect on the treated (ATT).

For nonnegative outcomes with a substantial mass at zero (weekly trips, total spending, and total items), we estimate Poisson pseudo-maximum likelihood (PPML) fixed-effects models. The PPML specification is well-suited to our setting. First, it naturally accommodates nonnegative outcomes, including the zeros arising in weeks when consumers did not shop, and unlike log-transformations of the outcome, does so without distorting effect magnitudes (Chen and Roth, 2024). Second, the estimator is consistent under correct specification of the conditional mean alone, without requiring the dependent variable to be a count or to follow a Poisson distribution, making it robust to overdispersion (Santos Silva and Tenreyro, 2006). The Poisson model coefficients can be interpreted as proportional changes using $\exp(\beta_2) - 1$ as the exact proportional change in the conditional mean per unit increase in treatment intensity. For per-trip and per-item averages (average price per item, spending per trip, items per trip) that do not contain any zeros, we use linear (OLS) fixed effects models with the same fixed effect structure. For the extensive margin (whether a consumer makes any trip in a given week), we estimate a logit specification. In all specifications, standard errors are clustered at the consumer level, the level at which treatment varies, allowing for arbitrary serial correlation within consumers over time (Bertrand et al., 2004; Abadie et al., 2023).

4.2 Identification

Our identification strategy relies on several key assumptions and design features.

Parallel trends. The central identifying assumption is that, absent CLS store adoption, treated and control consumers would have exhibited similar trends in shopping behavior. Several features of our design support this assumption. First, the inclusion of individual fixed effects (α_i) absorbs all time-invariant differences between consumers, including baseline shopping levels, geographic location, and demographic characteristics. Week fixed effects (γ_t) absorb common temporal shocks (such as seasonality, holidays, or macroeconomic conditions) affecting all consumers. Second, both treated and control consumers are drawn from the same pool of CLS adopters (as described in Section 3), and the coarsened exact matching on pre-treatment shopping behavior further supports comparability in baseline trends. Third, we provide direct evidence on pre-treatment balance in Table A1, which shows that the matched treated and control groups exhibit similar pre-adoption shopping patterns.

To evaluate this assumption further, we use event study analysis to examine how consumers change their behavior after adoption. We find that the pre-adoption spending and trip trends are statistically indistinguishable between the treatment and control groups. Full details of this analysis are provided in Section B of the Web Appendix.

Because our primary specification uses a continuous dose, the parallel-trends logic also extends to treatment intensity: absent treatment, consumers who went on to use CLS more heavily must have trended similarly to those who used it more lightly (Callaway et al., 2024). The event-study evidence supports this stronger requirement as well. When we split treated consumers into terciles of treatment intensity, all three groups display similar, flat pre-adoption trajectories (Figure B3 in the Web Appendix), indicating that heavier eventual usage is not associated with differential pre-trends.

Staggered adoption. Because adoption is staggered across consumers, standard two-way fixed effects (TWFE) estimators can be biased when treatment effects differ across adoption cohorts, as already-treated units may serve as implicit controls for later adopters (e.g., Sun and Abraham, 2021; Goodman-Bacon, 2021). In our design, the one-trip control group never enters a post-treatment state, so treated-versus-control comparisons are available in every calendar week and the TWFE estimator places relatively little weight on comparisons between earlier and later adopters. Such comparisons are not eliminated entirely, however. We therefore re-estimate a binary treated/control version of the design using the extended TWFE estimator of Wooldridge (2023) and the interaction-weighted estimator of Sun and Abraham (2021), both of which exclude already-treated units as controls and run within the same Poisson fixed-effects framework as our main specification. Both yield positive and statistically significant effects on weekly trips, consistent with our main results (Table C1 in Appendix C).

Within-adopter comparison. A distinctive feature of our design is that the control group consists of consumers who tried CLS stores but did not become regular users, rather than consumers who never visited a CLS store. This within-adopter comparison mitigates concerns about selection on unobservable technology preferences or awareness: both groups took the same initial step of entering a CLS store and completing a transaction. The relevant variation is in sustained usage, not in the initial adoption decision. To the extent that one-trip consumers experienced a small treatment from their single visit, our estimates are conservative—they capture the incremental effect of regular CLS usage relative to minimal exposure, rather than the effect relative to zero exposure. Consistent with the one-trip group remaining a valid benchmark after its single visit, the Post $\times$ Treated coefficient in our main specifications is small and statistically insignificant (Table 5): treated and control consumers show no post-adoption level divergence beyond the treatment-intensity effect itself.

Potential threats. The primary identification concern is that treatment intensity may be endogenous: consumers who increase their overall shopping frequency for unrelated reasons could mechanically accumulate more CLS trips. We address this threat in two ways. First, the continuous intensity measure is constructed using total post-period CLS trips divided by post-period weeks observed, so it reflects the average rate of CLS usage rather than a cumulative count that grows mechanically with time. Second, in Section K.1, we construct a shift-share instrumental variable based on pre-adoption geographic shopping patterns and the staggered rollout of CLS activity across districts, estimated on the subsample of geographically stable consumers for whom shopping locations meaningfully predict access. The IV independently supports the extensive-margin and basket-composition findings, and has consistent results with the results on weekly totals.

Exclusions. We exclude the adoption week itself (relative week zero) from all regressions to ensure clean separation of pre- and post-adoption periods and also because the adoption week mechanically includes a CLS store visit.

5 Results

5.1 Overall Shopping Behavior

We first examine how CLS usage is associated with each consumer's overall shopping behavior—their total weekly trips, spending, and items measured across the retailer's entire network of more than 9,000 stores, of which the vast majority are traditional staffed locations. These outcomes therefore capture total shopping across all formats, not activity at CLS stores alone. Table 5 presents Poisson regressions for these weekly totals. The key coefficient is the triple interaction of post-treatment status, treatment group, and treatment intensity, which captures how the treatment effect scales with the consumer's average weekly CLS trips. The standalone Post coefficient, by contrast, is not a treatment effect but a baseline pre-to-post term. It is statistically significant largely because adoptions are concentrated in the early-winter months, so on average it reflects seasonal (winter) changes in consumption rather than any consequence of CLS usage.

	Trips	Spending	Items
Post	0.051**	0.056*	0.059**
	(0.017)	(0.023)	(0.022)
Post:Treated	0.018	-0.016	-0.003
	(0.022)	(0.027)	(0.026)
Post:Treated:CLSUsage	0.238***	0.148***	0.195***
	(0.024)	(0.026)	(0.025)
Num.Obs.	165811	165811	165811
R^2	0.175	0.321	0.258
RMSE	1.96	36.89	7.25
FE: Consumer	X	X	X
FE: Week	X	X	X

Table 5: Treatment Effects on Weekly Shopping Totals. Poisson regressions with individual and week fixed effects. Standard errors, clustered by consumer, in parentheses. $^{+}\,p < 0.1$; $^{*}\,p < 0.05$; $^{**}\,p < 0.01$; $^{***}\,p < 0.001$.

Table 6 examines the intensive and extensive margins: average outcomes per trip and per item, and whether the consumer makes any trip in a given week.

	Price/Item	Spending/Trip	Items/Trip	Logit (Extensive)
Post	-0.021	-0.340*	-0.054+	0.055
	(0.040)	(0.161)	(0.029)	(0.036)
Post:Treated	-0.069	-0.187	0.019	-0.078+
	(0.044)	(0.193)	(0.037)	(0.047)
Post:Treated:CLSUsage	-0.406***	-1.749***	-0.219***	0.851***
	(0.058)	(0.256)	(0.051)	(0.080)
Num.Obs.	119532	119532	119532	159583
R^2	0.182	0.318	0.294	0.156
RMSE	2.60	10.55	1.90	0.41
FE: Consumer	X	X	X	X
FE: Week	X	X	X	X

Table 6: Treatment Effects on Shopping Intensity and Extensive Margin. OLS regressions for per-trip averages and a logit regression for the extensive margin (any trip in a given week), with individual and week fixed effects. Standard errors, clustered by consumer, in parentheses. $^{+}\,p < 0.1$; $^{*}\,p < 0.05$; $^{**}\,p < 0.01$; $^{***}\,p < 0.001$.

A one-unit increase in sustained CLS usage (one additional average CLS trip per week) is associated with approximately a 27% increase in weekly trip frequency, compared with increases of about 16% in spending and 22% in items purchased.[3] On the extensive margin, the same unit increase raises the probability of making any trip in a week by roughly 14 percentage points. At the trip level, spending falls by about 1.75 LCU, item count by about 0.22 items per trip, and average price per item by about 0.41 LCU. Thus, CLS usage is associated with some incremental spending but disproportionately increases trip incidence, fragmenting demand across a larger number of smaller baskets. Part of the per-trip decline is due to CLS trips being smaller than the consumer's typical trip. However, baskets at traditional stores also shrink as CLS usage rises (see Appendix E). This indicates that the shift toward smaller shopping occasions is not confined to the new format, but also occurs at traditional stores.

5.2 Category-Level Effects

Consumer responses to CLS store usage could vary across product types, reflecting both the direct effect of assortment differences and indirect effects of shifts in shopping occasions and trip patterns. Table 7 presents descriptive statistics on spending and purchase patterns across major product

[3] E.g., $\exp(0.238) - 1 \approx 27\%$ (Section 4.1).

categories. Beverages represent the largest category by both spending and items purchased, followed by Tobacco and Confectionery. Average price per item varies considerably, from 4 LCU for Beverages and Confectionery to 22 LCU for Strong Alcohols.

Rank	Category	% Spending	Avg. Price (LCU)	Unique Products	% Items	Items Rank
1	Beverages	24	4	722	29	1
2	Tobacco	12	13	195	4	8
3	Confectionery	9	4	773	11	2
4	Cafe	9	5	155	8	3
5	Beer	7	4	334	8	4
6	Grab and Go	6	7	196	4	9
7	Dairy	5	4	530	6	5
8	Strong Alcohols	5	22	305	1	16
9	Snacks	5	5	279	5	7
10	Grocery Food	3	4	827	3	10
11	Wine	3	19	212	1	19
12	Ready meals	2	8	102	1	15

Table 7: Descriptive Statistics by Product Category. Share of total spending and of total items, average price per item (LCU), and number of unique products, for the top 12 categories by total spending.

Figure 2 presents treatment effects separately for each major product category, with categories ordered by coefficient magnitude to facilitate comparison. We estimate our baseline specification (Equation 1) using separate Poisson regressions for each category-outcome combination. The panel is at the individual-week level, and the outcome is the consumer's weekly total of category-specific trips, spending, or items, with zero weeks retained as in Table 5. Each category regression uses the consumers who purchased from that category before their own adoption date, so the sample varies across categories.

In Figure 2, the largest increases in weekly spending are for grab-and-go food, dairy, confectionery, and beverages. Grocery food and cafe goods show smaller increases. This suggests that CLS stores are especially helpful for facilitating purchases of goods that are aligned with on-the-go consumption, rather than weekly grocery purchases or pantry staples.

There are also four restricted categories that are not available in CLS stores: beer, tobacco, wine, and strong alcohols. For these categories, Figure 2 shows declines in trips, spending, and items. The estimates suggest Strong Alcohols (i.e., spirits) as the category with the largest decline in weekly spending associated with CLS usage. Each additional CLS trip per week is associated with a 34% reduction in weekly spirits spending among spirits buyers ($p = 0.001$).[4] The beer spending decline is marginally significant (-13%, $p = 0.063$). Wine and tobacco show negative but

[4] E.g., $\exp(-0.412) - 1 \approx -34\%$.

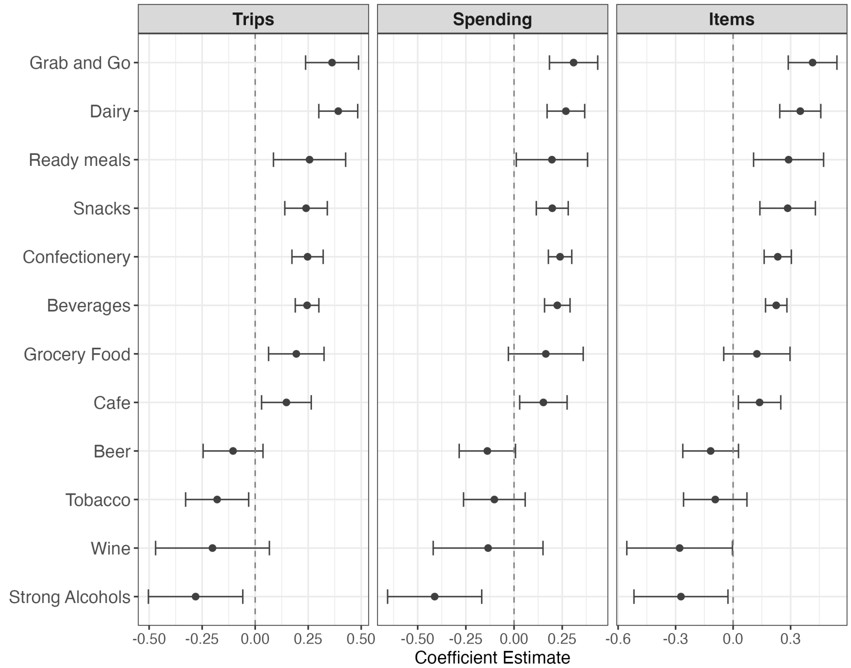

Figure 2: Treatment Effect Coefficients by Product Category. Each point shows the treatment intensity coefficient (Post × Treated × CLSUsage) from a separate Poisson regression of the consumer's weekly total category outcomes, with zero weeks retained, as in Table 5. Each regression uses the consumers with positive pre-adoption spending in that category. Error bars represent 95% confidence intervals from standard errors clustered by consumer.

insignificant spending declines of −13% and −10%. Section H of the Web Appendix reports these estimates in table form.

Even though these products are not available in CLS stores, CLS consumers can still buy them from traditional, non-CLS stores, which we continue to observe. Section 5.5 examines these restricted category declines in further detail.

5.3 Volumetric Analysis

Beyond counts and spending, we examine whether CLS store usage is associated with changes in the physical characteristics of items purchased. We extract weight (in grams) and volume (in milliliters) from product names in our transaction data. This allows us to assess whether consumers are purchasing larger or smaller package sizes, which may reflect changes in shopping trip purpose (e.g., stock-up trips vs. immediate consumption). 87% of purchased items carry at least one of these two measures: volume is identified for 48% of items (predominantly beverages) and weight for 39% (e.g., snacks and confectionery). We aggregate these measures to the individual-week level, retaining weeks with zero totals rather than excluding them.

Table 8 presents results for both weekly totals (Poisson, including zero-outcome weeks) and per-item averages (OLS, conditional on purchasing items with extractable measures). We find that CLS usage is linked with purchasing smaller items. In columns 3 and 4, we find a significant reduction in per-item weights and volumes among consumers who use CLS stores more.

	Weight	Volume	Weight/Item	Volume/Item
Post	0.064*	0.087*	-0.275	18.758
	(0.028)	(0.035)	(2.129)	(24.645)
Post:Treated	-0.008	-0.047	-3.576	-33.748
	(0.033)	(0.039)	(2.763)	(23.819)
Post:Treated:CLSUsage	0.229***	0.156***	-6.602*	-56.332***
	(0.035)	(0.035)	(2.929)	(15.035)
Num.Obs.	165744	165811	88577	102453
R^2	0.323	0.323	0.172	0.106
RMSE	560.38	4378.98	130.52	1323.84
FE: Consumer	X	X	X	X
FE: Week	X	X	X	X

Table 8: Treatment Effects on Product Weight (Grams) and Volume (Milliliters). Columns 1–2: Poisson regressions for weekly totals, including zero-outcome weeks. Columns 3–4: OLS for per-item averages, conditional on purchasing items with extractable measures. Individual and week fixed effects. Observation counts differ across columns: the totals drop consumers whose weekly total is always zero, while the per-item columns keep only weeks with a relevant purchase. Standard errors, clustered by consumer, in parentheses. $^+ p < 0.1$; $^* p < 0.05$; $^{**} p < 0.01$; $^{***} p < 0.001$.

5.4 Time-of-Day Analysis

Autonomous retail can also change when consumers shop. To examine changes in shopping timing, we partition the day into four time periods based on typical shopping patterns and store operations: Morning (6:00–10:00, representing 13% of trips across the data), before and during the start of the workday; Work Hours (10:00–16:00, 37%), during typical business hours; Evening (16:00–20:00, 31%), after work; and Night (20:00–23:00, 20%), before stores typically close.

The analysis uses the main-hours panel (6am–11pm) described in Section 3, which ensures a comparable observation window with traditional (non-CLS) stores. Trips outside this window account for only 0.7% of activity. Some CLS locations operate around the clock, and adopters do make more middle-of-the-night visits after adoption, but these remain a small share of total trips.

Empirical specification. To examine time-of-day effects, we extend our baseline Poisson model to include time period interactions. We construct a panel at the individual-week-time period level, where the outcome is the number of trips in each time period within each week:

$$\begin{aligned} Y_{itp} = {} & \alpha_i + \gamma_{tp} + \lambda_{ip} + \beta_0 \cdot \text{Post}_{it} + \beta_1 \cdot \text{Post}_{it} \times \text{Treated}_i + \beta_2 \cdot \text{Post}_{it} \times \text{Treated}_i \times \text{CLSUsage}_i \\ & + \sum_{p' \neq \text{Morning}} \delta^{p'} \cdot \text{Post}_{it} \times \text{Treated}_i \times \text{CLSUsage}_i \times \mathbf{1}[p = p'] + \varepsilon_{itp} \end{aligned} \tag{2}$$

where p indexes time periods, γ_{tp} are week-by-time-period fixed effects that absorb common calendar shifts in the timing of shopping across the day (for example, seasonal changes in evening activity), and λ_{ip} are individual-by-time-period fixed effects that control for each consumer's baseline propensity to shop at different times of day. The coefficient β_2 is the treatment-intensity coefficient for the Morning reference period, and the differential coefficients $\delta^{p'}$ for the Work, Evening, and Night periods capture how the effect in each of those periods differs from Morning.

We estimate this specification in two complementary ways. The count-based form in Equation 2 asks whether trips in each period rise with usage. The proportion-based form adds the log of total weekly trips, $\log(\text{TotalTrips}_{it})$, as an offset, so that the coefficients capture changes in the proportion of weekly trips in each time period, holding total trips constant. This shows if CLS usage shifts the distribution of trips across the day.

Table 9 presents the time-of-day regression results. Because the specification includes week-by-time-period fixed effects, common calendar shifts in the timing of shopping are absorbed, and the coefficients isolate the reallocation associated with CLS usage. Heavier usage is associated with a higher proportion of weekly trips taken in the morning, rather than an even scaling across all time periods. The morning coefficient in the proportions column is positive and significant (0.068, $p < 0.05$), the Work and Evening differentials relative to morning are negative (-0.090 and -0.081, both $p < 0.05$), and the Night differential is negative but imprecisely estimated (-0.045).

5.5 Restricted Categories

Tobacco and alcohol are not sold at CLS stores, and Figure 2 showed declines in these restricted categories associated with CLS usage. We examine this result in further detail among the consumers who previously purchased restricted categories in the pre-adoption period. Restricted purchases are part of these consumers' routine shopping. Before adoption, buyers average 2.3 trips per week, of

	Totals	Proportions
Post	0.049**	-0.001
	(0.017)	(0.001)
Post:Treated	0.017	-0.001
	(0.022)	(0.001)
Post:Treated:CLSUsage	0.317***	0.068*
	(0.043)	(0.035)
Post:Treated:CLSUsage:Work	-0.102*	-0.090*
	(0.046)	(0.042)
Post:Treated:CLSUsage:Evening	-0.096*	-0.081*
	(0.043)	(0.041)
Post:Treated:CLSUsage:Night	-0.046	-0.045
	(0.050)	(0.046)
Num.Obs.	642955	466371
RMSE	0.87	0.76
FE: Consumer	X	X
FE: Week × time period	X	X
FE: Consumer × time period	X	X

Table 9: Treatment Effects on Shopping Trips by Time of Day. Poisson regressions with individual, week-by-time-period, and individual-by-time-period fixed effects. Morning is the reference period; coefficients for other periods are differentials relative to Morning. Standard errors, clustered by consumer, in parentheses. $^{+}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

which 0.6—about one in four—include a restricted item. Specifically, we consider if the observed declines appear more on the extensive or intensive margins.

Table 10 reports six outcomes among the pooled restricted categories (beer, wine, tobacco, and strong alcohol). Each additional weekly CLS trip is associated with about 15% fewer weekly restricted trips and 15% less weekly restricted spending. The reduction appears to be due to fewer purchases rather than smaller ones. Column 3 shows the probability of any restricted purchase in a week falls; Columns 4 and 5 show that, even within purchase weeks, restricted trips fall by about 7% and spending by about 8%. Spending per restricted basket, on the other hand, has no significant change. Section 7.1 further examine how these declines vary with consumer types.

	Trips	Spending	Any Purchase	Trips (Purchase Weeks)	Spend (Purchase Weeks)	Spend per Basket
Post	0.020	0.048	-0.045	0.033	0.053+	-0.013
	(0.034)	(0.039)	(0.040)	(0.021)	(0.028)	(0.019)
Post:Treated	0.014	-0.017	-0.016	0.033	0.019	-0.015
	(0.043)	(0.047)	(0.052)	(0.024)	(0.031)	(0.027)
Post:Treated:CLSUsage	-0.168**	-0.165**	-0.126+	-0.071*	-0.087*	-0.029
	(0.054)	(0.057)	(0.076)	(0.030)	(0.042)	(0.047)
Num.Obs.	138889	138889	138796	41489	41489	41489
R^2	0.249	0.346	0.214	0.089	0.296	0.233
RMSE	1.01	21.23	0.40	1.30	30.43	12.76
FE: Consumer	X	X	X	X	X	X
FE: Week	X	X	X	X	X	X

Table 10: Restricted-Category Outcomes Among Buyers. Sample: consumers with positive pre-treatment restricted spending. Columns 1–2: Poisson on weekly restricted trips and spending, including zero weeks. Column 3: logit on any restricted purchase in the week. Columns 4–5: Poisson conditional on a purchase week. Column 6: Poisson on spending per restricted basket, conditional on a restricted trip. Individual and week fixed effects. Standard errors, clustered by consumer, in parentheses. $^+ p < 0.1$; $^* p < 0.05$; $^{**} p < 0.01$; $^{***} p < 0.001$.

6 Mechanisms

The preceding section presents three core findings. First, CLS store adopters shop more frequently, with total spending rising substantially less than trip frequency, so purchases fragment into smaller baskets. Second, spending in some restricted categories declines. Third, the trip-frequency increase is concentrated in morning hours.

We first relate these patterns to the time cost of shopping and examine trip fragmentation directly. We then posit three mechanisms that could produce the finding of smaller per-item weights and volumes: (a) CLS stores may carry a narrower assortment that excludes larger package sizes (assortment constraint); (b) consumers may generally prefer smaller sizes across all trips when shopping more frequently (inventory response); or (c) CLS trips may constitute a different *type* of shopping occasion that calls for grab-and-go sizes (occasion mechanism). These mechanisms are not mutually exclusive. The analyses below examine the support for each.

6.1 Time Pressure and Frictionless Retail

The concentration of the treatment effect in morning hours is consistent with a time-pressure mechanism. The value of a fast, predictable shopping trip may be greatest when the consumer's opportunity cost of time is high.

Cashierless CLS stores reduce this fixed cost. By eliminating the checkout process entirely, a CLS trip takes little more than the time needed to pick up the items. The time savings may be modest in absolute terms, but during high-time-pressure periods, even small reductions can shift a shopping occasion from skipped to taken. This is consistent with the wait-time sensitivity documented by Allon et al. (2011) and with the broader literature on how time costs shape shopping behavior (Ülkü et al., 2020).

The differential effect across time periods is consistent with this interpretation. The analysis in Section 5.4 shows that heavier CLS usage is associated with a higher proportion of trips taken in the morning rather than an even scaling of trips across the day. The disproportionate morning shift provides suggestive evidence that eliminating checkout friction is particularly valuable when the opportunity cost of time is high. Because CLS store technology enables placement in locations that were previously difficult to access, however, we interpret this as evidence consistent with a checkout-friction mechanism rather than a design that separately identifies it.

6.2 Inter-Purchase Time

The introduction of CLS stores may encourage consumers to split up big shopping trips into smaller, more frequent trips. We presented initial evidence in Section 5.1, but to examine this further, we calculate the inter-purchase time (IPT) between trips and find that IPT among CLS adopters is about 2 days shorter on average, relative to the control group. Details of this analysis are in Section D of the Web Appendix.

6.3 Assortment and Prices

Since CLS stores typically have fewer products than traditional (non-CLS) stores, treated consumers who visit CLS stores might be seeing a smaller assortment of products compared to consumers in the control group. We examine this issue in Section E of the Web Appendix by estimating treatment effects on weekly store-choice outcomes. Two results show that CLS trips add to consumers' existing stores rather than replace them. First, each additional weekly CLS trip is associated with visiting 0.14 more distinct stores per week ($p < 0.001$). Second, the number of distinct traditional stores visited is essentially unchanged (-0.04, $p > 0.1$). Consumers appear to keep visiting the same number of traditional stores they used before adopting CLS. Therefore, we assume that treated consumers retain regular access to the full assortment sold at traditional stores, which weakens the assortment-constraint explanation (a).

A related question is whether treated consumers see a very different set of prices than control group consumers. This might happen if the retail chain charges different prices for the same products in CLS vs. traditional (non-CLS) stores. We compare each product's mean unit price at CLS stores with its mean unit price at traditional stores, for products sold at both formats (see Section F of the Web Appendix). On average, CLS prices are 0.69% lower, not statistically different from zero. In comparison, this is an order of magnitude smaller than existing median 8.6% price dispersion across traditional stores. Price differences are therefore an unlikely driver of the behavioral changes we observe.

6.4 Volume Decline: Composition vs. Size Choice

In Section 5.3, we find that CLS adopters purchase items with lower per-unit weight and volume. We investigate two explanations for this outcome. First, there might be between-product effects where CLS adopters change which categories and brands they are buying. Second, there might be within-product effects where CLS adopters buy smaller sizes of the same type of base product.

As a back-of-the-envelope analysis, we conduct a within-individual decomposition of pre- to post-treatment changes in product choices in Section G of the Web Appendix. First, for all volume products, we identify the *base product* independent of specific sizes, such as Pepsi 8 ounce versus Pepsi 2 liter. Then, for each consumer, we observe their change in average package volume (mL per item). This can be decomposed into changes in size versus changes in base product composition. We calculate size choice by holding each consumer's pre-treatment base product mix constant, and measure how much the average changes because the consumer chooses smaller versions of the same base product. The composition becomes the remainder, measuring how much change is because the consumer shifted their product mix to different base products (with each base measured in the post-period average size).

Averaging across treated consumers, the average package falls by 6.1 mL per item. Within-product size choice accounts for about a third (−1.7 mL) while composition accounts for about two-thirds of the decline (−4.4 mL). This only considers base products bought in both periods.

While that is a back-of-the-envelope analysis, we conduct a specific regression test in Section G of the Web Appendix as well. We identify consumers who buy the same base volume product in both CLS and traditional stores in the post-period. In this analysis, we find that they choose a package that is on average 142 mL smaller at a CLS store than at a traditional store ($p < 0.001$). This estimate remains similar when we further restrict the sample to products available in multiple sizes at CLS stores. This suggests that consumers may be preferring to choose smaller sized items at CLS stores, even when larger sizes are available at the same time.

Lastly, we analyze baskets at traditional stores for CLS adopters in Section E of the Web Appendix. We find that CLS usage is associated with reduced trips to traditional stores, and in these traditional store visits, a significant reduction in the number of items per basket (0.1, $p < 0.05$). These traditional stores are not significantly different in assortment though, so they are not smaller traditional stores. We also find no significant change in the sizes of items purchased associated with CLS usage (Appendix Table E3). CLS usage is associated with a directionally negative weight and volume per item, but insignificantly so.

Taken together, we find that CLS stores are not priced differently, and that consumers appear to have the same access to traditional assortments between the pre- and post- periods. CLS usage is associated with evidence of voluntary product downsizing, with smaller product choices at CLS stores even when multiple sizes are available. This downsizing is not observed significantly at traditional stores however. An inventory mechanism may explain slightly smaller baskets at traditional stores, but the size-choice findings are concentrated in CLS stores, consistent with CLS enabling a new shopping occasion.

7 Consumer Heterogeneity

The average treatment effects may conceal meaningful variation across consumer types. We examine heterogeneity along three consumer dimensions: levels of alcohol and tobacco (restricted-category) spending, shopping style, and time-of-day shopping profiles. All consumer segments are defined using only pre-treatment data to avoid conditioning on post-treatment outcomes.

7.1 Heterogeneity by Pre-Treatment Restricted-Category Spending

In the overall data, we find a significant reduction in trips that involve strong alcohols and tobacco (and marginally significant reduction in trips with beer and wine). It is a priori uncertain whether this is associated more with light-spenders in these restricted categories, or heavy spenders.

We examine how the decline in these restricted-category purchases varies with consumers' pre-treatment history of buying these items. Among consumers with non-zero pre-treatment restricted spending, we split below vs. above the median into "Light" and "Heavy" groups. In the pre-treatment period, the Light group averaged 3.2 LCU per week in restricted categories and the Heavy group averaged 16.5 LCU per week. The two groups shop similarly often (about 2.3 trips per week) but they differ in how many of those trips carry restricted items (10% versus 43%).

We estimate the same specification as Equation 1, separately for the Light and Heavy groups, with weekly restricted-category trips and spending as outcomes. Because each group is defined by a pre-period median split, the Post coefficient absorbs the group's common pre-to-post shift, which combines winter seasonality with any reversion toward the mean we may observe due to the group splitting. Table 11 reports the estimates.

	Trips (Light)	Trips (Heavy)	Spend (Light)	Spend (Heavy)
Post	0.149*	-0.012	0.301***	-0.008
	(0.061)	(0.039)	(0.068)	(0.045)
Post:Treated	0.086	-0.009	0.069	-0.042
	(0.081)	(0.051)	(0.090)	(0.055)
Post:Treated:CLSUsage	-0.181+	-0.164*	-0.179	-0.167**
	(0.094)	(0.064)	(0.120)	(0.063)
Num.Obs.	70023	68866	70023	68866
R^2	0.165	0.186	0.230	0.259
RMSE	0.60	1.30	11.32	27.88
FE: Consumer	X	X	X	X
FE: Week	X	X	X	X

Table 11: Treatment Heterogeneity: Restricted Categories by Pre-Treatment Restricted Spending. Poisson regressions for weekly restricted-category trips and spending, estimated separately for the Light and Heavy segments (below/above-median pre-treatment restricted spending among buyers). Standard errors, clustered by consumer, in parentheses. $^{+}\,p < 0.1$; $^{*}\,p < 0.05$; $^{**}\,p < 0.01$; $^{***}\,p < 0.001$.

The proportional effects are similar for the two groups. Each additional weekly CLS trip is associated with roughly 15–17% less restricted spending and fewer restricted-item trips for both Light and Heavy buyers.[5] The estimates are precise for the Heavy group (spending $p < 0.01$, trips $p < 0.05$) and weaker for the Light group (trips $p = 0.054$; spending not significant).

Given the two groups' differing baseline spending amounts, the absolute LCU interpretation differs between the two groups. The Heavy group is associated with roughly 15% lower weekly

[5] E.g., $\exp(-0.167) - 1 \approx -15\%$.

spending on restricted categories per weekly CLS trip, which at their pre-adoption baseline is about 2.5 LCU per week.

7.1.1 Light and Heavy Buyers by Restricted Category

We repeat the Light and Heavy split for each restricted category separately, classifying buyers by their pre-adoption spending in that category (Section H of the Web Appendix). We discuss Strong Alcohols here, the category with the largest spending decline in Figure 2. Light and Heavy spirits buyers respond similarly in proportional terms (-36% and -31%). Each additional weekly CLS trip removes roughly a third of spirits spending for both groups. The Heavy-spirits buyers average 6.34 LCU per week on spirits across their pre-adoption panel weeks. At their -31% response, a Heavy-spirits buyer averaging one CLS trip per week is estimated to reduce spirits spending by about 2 LCU per week.

7.2 Shopping-Style and Time-of-Day Segments

We utilize data-driven clustering of consumers based on their (a) shopping styles and (b) time-of-day profiles.

7.2.1 Shopping-Style Clusters

We cluster consumers based on a few characteristics of how they shop. We measure pre-treatment behavioral features including weekly trip frequency, average basket size (items and LCU per trip), and across the pre-treatment period, the total number of distinct stores visited and the median days between trips. K-means clustering on these standardized features yields three groups. We characterize these groups as *Style 1: light shoppers* (2.1 trips/week, 23 LCU/week, 2.6 items/trip, 10 distinct stores; n = 2,375), *Style 2: big-basket shoppers* (2.5 trips/week, 64 LCU/week, 5.0 items/trip, 10 distinct stores; n = 734), and *Style 3: frequent multi-store shoppers* (4.0 trips/week, 50 LCU/week, 2.7 items/trip, 22 distinct stores; n = 1,329). We replicate Equation 1 for each of the three subgroups. Table 12 reports the estimates.

We observe that light and big-basket shoppers have approximately twice as strong changes in trips and spending associated with CLS usage, compared to the frequent multi-store shoppers. This is consistent with shopping frequency being an important factor. Simple terciles of pre-treatment trip frequency show the same declining gradient (Section I of the Web Appendix). Consumers who already shop frequently across many stores may have less room for CLS to add value, whereas light and/or less-frequent shoppers may react to CLS as a top-up channel that opens new shopping occasions.

	Trips (Style 1)	Trips (Style 2)	Trips (Style 3)	Spend (Style 1)	Spend (Style 2)	Spend (Style 3)
Post	0.126***	0.109*	-0.027	0.177***	0.035	-0.015
	(0.025)	(0.044)	(0.026)	(0.031)	(0.056)	(0.035)
Post:Treated	0.036	-0.082	0.031	0.041	-0.139*	0.005
	(0.031)	(0.050)	(0.034)	(0.037)	(0.056)	(0.045)
Post:Treated:CLSUsage	0.324***	0.299***	0.149***	0.202***	0.203***	0.084*
	(0.033)	(0.048)	(0.031)	(0.036)	(0.047)	(0.041)
Num.Obs.	86312	26849	52650	86312	26849	52650
R^2	0.107	0.145	0.113	0.190	0.246	0.237
RMSE	1.49	1.76	2.60	23.59	53.95	43.16
FE: Consumer	X	X	X	X	X	X
FE: Week	X	X	X	X	X	X

Table 12: Treatment Heterogeneity by Pre-Treatment Shopping Style. Poisson regressions estimated separately by shopping-style cluster (k-means on pre-treatment behavioral features; see text). Individual and week fixed effects. Standard errors, clustered by consumer, in parentheses. $^{+}\,p < 0.1$; $^{*}\,p < 0.05$; $^{**}\,p < 0.01$; $^{***}\,p < 0.001$.

7.2.2 Hourly Shopping Profiles

We also cluster consumers based on their distribution of shopping times within a day. For consumers with at least five pre-adoption trips, we calculate the percent of trips in each one-hour bin from 6am to 11pm (k-means, $k = 4$). The clustering largely reflects the four named time periods, with one cluster peaking in each. There is a morning cluster (peak 7–8am; 42% of its trips in the morning window; $n = 424$), a work/midday cluster (peak 1–2pm; 54% in the work window; $n =$ 1,384), an evening cluster (peak 5–6pm; 45% in the evening window; $n =$ 1,260), and a night cluster (peak 8–9pm; 37% in the night window; $n =$ 1,310).

Estimated separately within each cluster (Figure 3), the treatment-intensity coefficients are consistent with an approximate U-shape over the day. The morning cluster has the largest change associated with CLS usage. The response is lower for the work/midday and evening clusters, and the point estimate is slightly larger for the night cluster. A $k = 5$ cluster variation in Section J of the Web Appendix shows a similar U-shape over consumers' preferred shopping times. This U-shape is consistent with the time-pressure interpretation of Section 6.1. We treat the cross-cluster comparisons as descriptive, however.

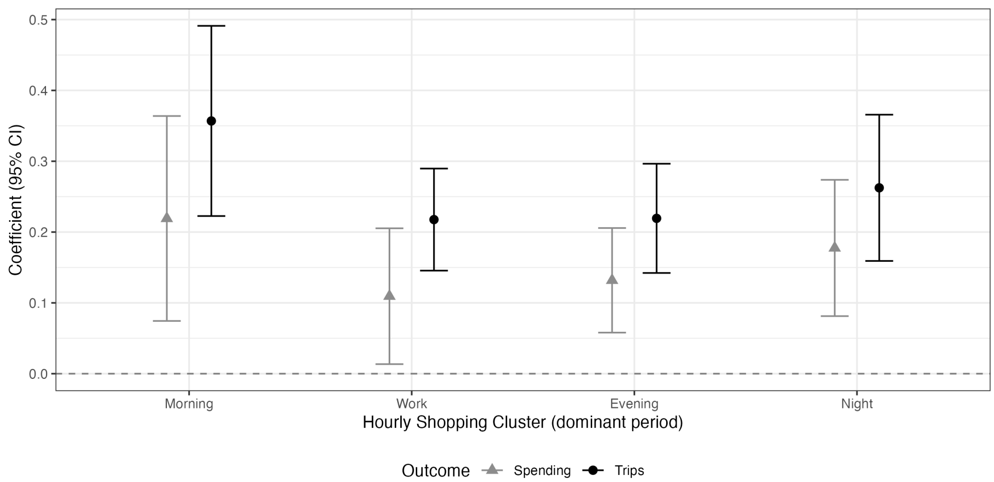

Figure 3: Treatment heterogeneity by hourly shopping profile (exploratory). Consumers are clustered (k-means, $k = 4$) on their pre-adoption share of trips in each one-hour bin from 6am to 11pm; clusters are labeled by dominant time period. Points show the treatment-intensity coefficient (Post × Treated × CLSUsage) for trips and spending, with 95% confidence intervals. Table J1 in the Web Appendix presents the full regression coefficients underlying this figure.

8 Discussion

Cashierless retail is often described as a technology that removes checkout. Our findings suggest a broader effect: removing checkout can change the shopping occasion itself. In our setting, consumers who incorporate CLS stores into their routines shop more often, but in smaller baskets; they shift trips toward morning hours; they select smaller packages for the same products; and they reduce purchases in categories that are absent from the CLS format—most sharply spirits, where each additional weekly CLS trip is associated with roughly a third less spending among spirits buyers. These patterns indicate that cashierless stores are not simply faster versions of traditional stores. They create a distinct convenience occasion that allows consumers to purchase closer to the moment of need.

This conclusion is especially important for grocery retail. Grocery remains both economically large and operationally difficult for traditional e-commerce. Delivery-based grocery can solve some consumer convenience problems, but it also introduces costly fulfillment tasks such as picking, staging, substitution, and last-mile delivery. Cashierless CLS stores offer a different path: they preserve the immediacy and locality of the physical store while digitizing and automating the transaction. In this sense, the format is neither a substitute for supermarkets nor a smaller version of e-commerce. It is a distinct grocery format designed for small, fast, immediate, and time-sensitive trips.

This interpretation also helps explain why the industry evidence around cashierless retail has been mixed. The technology may be less compelling as a universal replacement for full-service grocery stores, where consumers buy large baskets, compare prices, seek assortment depth, and often want visibility into total basket cost. It may be more compelling in small-format environments where speed, predictability, and location accessibility dominate: office districts, hospitals, transit hubs, campuses, stadiums, residential complexes, and other locations where a staffed store may be difficult to operate profitably (Kumar, 2024; Briggs, 2023; Bensinger, 2024). Our findings are consistent with this view. The strongest behavioral effects occur in shopping contexts where the fixed cost of time is likely to matter most.

8.1 Theoretical implications

The first theoretical implication concerns self-service and AI-enabled retail technology. Prior research often treats self-service technology as a change in the service interface: consumers perform tasks that employees previously performed, and the main outcomes are adoption, satisfaction, perceived control, or service efficiency (Bulmer et al., 2018; Meuter et al., 2000; Nusrat and Huang, 2024). Cashierless retail pushes beyond this framing. In our setting, the technology does not ask consumers to do more work; it makes a stage of the shopping process disappear from the consumer's perspective. When the transaction process becomes sufficiently invisible, the consumer's decision shifts from choosing a checkout mode to deciding whether to make the trip at all. Our results show that technology-enabled friction reduction can change the incidence, timing, and content of shopping trips.

Second, the paper contributes to omnichannel and store-format research by showing that channel introductions can reconfigure demand even when they do not dramatically expand it. The standard expansion-versus-cannibalization framing is useful but incomplete (Avery et al., 2012; Bell et al., 2018; Neslin, 2022; Shriver and Bollinger, 2022). CLS stores do generate more trips, yet those trips are smaller and partly substitute for traditional-store routines. The relevant outcome is how consumers reorganize purchasing when they can buy in smaller, more immediate, and more time-constrained shopping occasions. This suggests that future omnichannel research should consider the shopping occasion as an intermediate mechanism between channel availability and demand outcomes.

Third, the paper contributes to research on shopping friction and time costs. The morning effect provides suggestive field evidence that the value of eliminating checkout friction is heterogeneous across time. A technology that saves a few minutes may have little effect during low-pressure periods but a large effect when consumers face binding schedule constraints (Allon et al., 2011; Ülkü et al., 2020). This implies that the same retail innovation can have different behavioral effects across dayparts, locations, and customer routines.

Fourth, the paper contributes to grocery shopping research by showing how technology can shift consumers along the stock-up-to-immediate-consumption continuum. Grocery shopping comprises several distinct missions. Consumers sometimes buy for the week, sometimes fill in missing items, and sometimes satisfy immediate consumption needs. Cashierless CLS stores appear to move some purchases toward the latter end of this continuum. The package-size results make this concrete: shopping under a quick-consumption occasion, consumers choose smaller quantities even of the same products. This is theoretically important because it shows that a store-format innovation can change the unit of analysis from the trip to the occasion; the channel can help define the basket.

Finally, the restricted-category results suggest that retail formats can shape habitual purchases by changing the context in which those purchases occur. Alcohol and tobacco are not available in CLS stores, and for some categories, most notably strong alcohol, consumers do not fully compensate by shifting those purchases to traditional stores. This does not mean cashierless stores are a public-health intervention, and we do not make such a broad claim. However, it does show that format design can disrupt habitual purchase routines, especially for categories that depend on cues, availability, and established shopping scripts. This finding connects to work showing that purchase environments can influence vice consumption, impulse buying, and socially sensitive purchases (Huyghe et al., 2017; Iyer et al., 2020; Massara et al., 2014; Sun et al., 2023).

8.2 Managerial implications

For retailers, the first implication is that cashierless stores should not be evaluated only by trip counts. More trips are not the same as proportionally more spending or profit. In our data, CLS usage is associated with more frequent but smaller, lower-priced trips. Retailers should therefore evaluate cashierless stores using a format-specific profit model that accounts for basket size, category mix, labor savings, replenishment costs, technology costs, shrinkage, and cannibalization from traditional stores. A successful CLS store may look weak if judged by average basket value, but strong if judged by throughput, labor productivity, customer retention, and access to otherwise unserved locations.

Second, location strategy is central. The format appears especially well suited to environments where time pressure is high and traditional staffing is costly or impractical—the transit, workplace, and event settings discussed above. These are places where consumers may have small immediate needs but little tolerance for waiting. In such settings, a cashierless store can convert a previously infeasible shopping occasion into a feasible one. By contrast, locations that mainly serve relaxed stock-up trips may not benefit as much from full cashierless automation.

Third, assortment strategy should reflect the occasion. CLS stores should not simply miniaturize a traditional convenience-store assortment. The evidence points toward quick, immediate-consumption missions: beverages, coffee, breakfast items, ready-to-eat foods, snacks, and small essentials. The morning shift suggests particular importance for breakfast and commute-oriented assortments. The smaller-package findings suggest that single-serve and portable formats may be especially valuable. Retailers should also recognize that the absence of high-margin restricted categories, such as tobacco and alcohol, can materially change the margin profile of cashierless stores.

Fourth, cashierless formats require different replenishment logic. Smaller baskets and morning concentration can create demand peaks even if total demand does not rise dramatically. Stores designed around quick trips may need more frequent replenishment of a narrower set of high-velocity items. Stockouts may be especially costly because the value proposition is convenience and predictability. If consumers enter a CLS store expecting a fast solution and the product is unavailable, the format's advantage weakens quickly.

Fifth, retailers should be clear about the strategic role of the format. If the goal is demand expansion, CLS stores should be placed where they attract new consumers or serve occasions not currently captured by the retailer. If the goal is retention, CLS stores may be valuable even when they cannibalize traditional stores, because they keep convenience-oriented occasions inside the retailer's ecosystem. If the goal is cost efficiency, the relevant comparison is not revenue per trip alone but contribution after labor, real estate, shrinkage, technology, and replenishment costs. The same technology can therefore be justified by different strategic logics in different locations.

The grocery-specific implication is that cashierless stores should not be positioned as a direct substitute for online grocery delivery. Online grocery is often strongest for planned, larger-basket shopping occasions. CLS stores are more naturally suited to small, immediate, high-frequency occasions where delivery economics are unfavorable and conventional stores are too slow or inconvenient. Retailers should therefore view CLS stores as a complement to, rather than a replacement for, both traditional stores and digital grocery channels. The strategic question is not "Should the future of grocery be online or offline?" but "Which grocery occasions should be served by which format?"

8.3 Policy and societal implications

The rise of cashierless stores also raises policy issues. On the positive side, small unstaffed formats may improve access in locations that cannot support a fully staffed store. They may provide late-night or early-morning access, serve transit and workplace nodes, and extend the reach of grocery and convenience retail. At the same time, the format raises questions about privacy, surveillance, algorithmic accuracy, accessibility for consumers without smartphones or payment cards, and employment effects (Benoit et al., 2024; De Bellis and Johar, 2020; Hazée et al., 2025; Grewal et al., 2023). These issues are not central to our empirical design, but they are important for understanding the broader deployment of cashierless retail.

The restricted-category results are also policy-relevant. In our setting, CLS stores do not sell alcohol and tobacco, and adoption is associated with reduced spending in some of these categories—most clearly strong alcohol, where the decline appears among both light and heavy pre-treatment buyers. This suggests that format design can influence the path of purchases by changing routine access and purchase cues. Policymakers should not infer that cashierless stores will generally reduce unhealthy purchases; effects will depend on assortment, regulation, and consumer substitution. Still, the finding illustrates how retail format innovation can have unintended behavioral consequences beyond operational efficiency.

8.4 Limitations and future research

In Section K of the Web Appendix, we conduct two robustness checks to examine how our key results change under alternative assumptions and specifications. First, we use a Bartik-style shift-share instrumental variable, estimated on the subsample of geographically stable consumers, to deal with potential endogeneity of treatment intensity. Second, we change the control group in our analysis to consumers who never visit CLS stores, rather than the one-trip controls. The alternative control group replicates the main results closely. The instrumental-variables analysis supports two of the paper's central findings using variation in access rather than choice: consumers with greater CLS access shop in more weeks and buy smaller, lower-priced baskets—the two margins at the center of the shopping-occasion mechanism in Section 6. The IV is too imprecise on its restricted subsample to assess the increases in total weekly trips, spending, and items; those magnitudes are identified from the comparison of heavier and lighter users within the full matched adopter sample.

Beyond these robustness checks, there are several limitations worth emphasizing. First, our estimates capture the effect of sustained CLS usage among adopters, not the average effect of randomly assigning all consumers to cashierless stores. This adopter-focused estimand is managerially relevant because retailers typically cannot compel consumers to choose a new store format, but it should not be interpreted as the effect of universal adoption. Second, the treatment combines technology and location. CLS stores are cashierless and placed in locations made viable by the cashierless format. Future research could use settings that separately vary checkout technology and store location. Third, we observe sales but not full profitability, technology costs, shrinkage, labor substitution, or competitor purchases. Future work could integrate cost data and competitive panels to evaluate the full economic returns of cashierless stores.

Despite these limitations, the study offers an important early view of how cashierless retail affects actual consumer behavior. The central lesson is that cashierless technology is best understood as an enabler of new retail occasions rather than a universal replacement for checkout. This is particularly true in grocery, where physical stores remain central because many purchases are immediate, local, and difficult to fulfill profitably through delivery. When shopping becomes sufficiently fast and predictable, consumers reorganize when they shop, how often they shop, and how much they buy at a time. This shift from faster transactions to different shopping occasions is the paper's broader implication for retail strategy.

References

Abadie, A., Athey, S., Imbens, G. W., and Wooldridge, J. M. (2023). When should you adjust standard errors for clustering? *The Quarterly Journal of Economics*, 138(1):1–35.

Allon, G., Federgruen, A., and Pierson, M. (2011). How much is a reduction of your customers' wait worth? An empirical study of the fast-food drive-thru industry based on structural estimation methods. *Manufacturing & Service Operations Management*, 13(4):489–507.

Amazon Web Services (2026). Amazon just walk out: Streamline operations and delight customers with intelligent, autonomous retail solutions. `https://aws.amazon.com/just-walk-out/`. Accessed: June 22, 2026.

Avery, J., Steenburgh, T. J., Deighton, J., and Caravella, M. (2012). Adding bricks to clicks: Predicting the patterns of cross-channel elasticities over time. *Journal of Marketing*, 76(3):96–111.

Aydinli, A., Lamey, L., Millet, K., ter Braak, A., and Vuegen, M. (2021). How do customers alter their basket composition when they perceive the retail store to be crowded? An empirical study. *Journal of Retailing*, 97(2):207–216.

Belavina, E., Girotra, K., and Kabra, A. (2017). Online grocery retail: Revenue models and environmental impact. *Management Science*, 63(6):1781–1799.

Bell, D. R., Gallino, S., and Moreno, A. (2015). Showrooms and information provision in omni-channel retail. *Production and Operations Management*, 24(3):360–362.

Bell, D. R., Gallino, S., and Moreno, A. (2018). Offline showrooms in omnichannel retail: Demand and operational benefits. *Management Science*, 64(4):1629–1651.

Benoit, S., Altrichter, B., Grewal, D., and Ahlbom, C.-P. (2024). Autonomous stores: How levels of in-store automation affect store patronage. *Journal of Retailing*, 100(2):217–238.

Benoit, S., Evanschitzky, H., and Teller, C. (2019). Retail format selection in on-the-go shopping situations. *Journal of Business Research*, 100:268–278.

Bensinger, G. (2024). Amazon to push cashierless shopping tech into more third-party stores, while backing off itself. `https://www.reuters.com/business/retail-consumer/amazon-push-cashierless-shopping-tech-into-more-third-party-stores-while-backing-2024-04-17/`. Accessed: June 22, 2026.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics*, 119(1):249–275.

Briggs, B. (2023). A grocer that sells smoothies, snacks and 'easier lives'? Microsoft Source News. Accessed: June 22, 2026.

Bulmer, S., Elms, J., and Moore, S. (2018). Exploring the adoption of self-service checkouts and the associated social obligations of shopping practices. *Journal of Retailing and Consumer Services*, 42:107–116.

Callaway, B., Goodman-Bacon, A., and Sant'Anna, P. H. C. (2024). Difference-in-differences with a continuous treatment. NBER Working Paper 32117, National Bureau of Economic Research.

Campbell, D. and Frei, F. (2010). Cost structure, customer profitability, and retention implications of self-service distribution channels: Evidence from customer behavior in an online banking channel. *Management Science*, 56(1):4–24.

Chen, J. and Roth, J. (2024). Logs with zeros? Some problems and solutions. *The Quarterly Journal of Economics*, 139(2):891–936.

Chintala, S. C., Liaukonytė, J., and Yang, N. (2024). Browsing the aisles or browsing the app? How online grocery shopping is changing what we buy. *Marketing Science*, 43(3):506–522.

Cui, Y., van Esch, P., and Jain, S. P. (2022). Just walk out: The effect of AI-enabled checkouts. *European Journal of Marketing*, 56(6):1650–1683.

De Bellis, E. and Johar, G. V. (2020). Autonomous shopping systems: Identifying and overcoming barriers to consumer adoption. *Journal of Retailing*, 96(1):74–87.

Fisher, M. L., Gallino, S., and Xu, J. J. (2019). The value of rapid delivery in omnichannel retailing. *Journal of Marketing Research*, 56(5):732–748.

FMI (2026). Food industry facts. `https://www.fmi.org/our-research/food-industry-facts`. Accessed: June 22, 2026.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277.

Grewal, D., Benoit, S., Noble, S. M., Guha, A., Ahlbom, C.-P., and Nordfält, J. (2023). Leveraging in-store technology and AI: Increasing customer and employee efficiency and enhancing their experiences. *Journal of Retailing*, 99(4):487–504.

Hazée, S., De Keyser, A., Larivière, B., Kim, Y., Talebi, A., and Gordeliy, I. (2025). Just walk out stores—the future of shopping? Examining configurations of reasons for and against consumer adoption. *Psychology & Marketing*, 42(4):987–1017.

Huyghe, E., Verstraeten, J., Geuens, M., and Van Kerckhove, A. (2017). Clicks as a healthy alternative to bricks: How online grocery shopping reduces vice purchases. *Journal of Marketing Research*, 54(1):61–74.

Iyer, G. R., Blut, M., Xiao, S. H., and Grewal, D. (2020). Impulse buying: A meta-analytic review. *Journal of the academy of marketing science*, 48(3):384–404.

Jindal, P., Zhu, T., Chintagunta, P., and Dhar, S. (2020). Marketing-mix response across retail formats: The role of shopping trip types. *Journal of Marketing*, 84(2):114–132.

Kumar, D. (2024). An update on Amazon's plans for just walk out and checkout-free technology. `https://www.aboutamazon.com/news/retail/amazon-just-walk-out-dash-cart-grocery-shopping-checkout-stores`. Accessed: June 22, 2026.

Massara, F., Melara, R. D., and Liu, S. S. (2014). Impulse versus opportunistic purchasing during a grocery shopping experience. *Marketing letters*, 25(4):361–372.

McKinsey & Company (2026). Achieving profitable online grocery order fulfillment. `https://www.mckinsey.com/industries/retail/our-insights/achieving-profitable-online-grocery-order-fulfillment`. Accessed: June 22, 2026.

Meuter, M. L., Ostrom, A. L., Roundtree, R. I., and Bitner, M. J. (2000). Self-service technologies: Understanding customer satisfaction with technology-based service encounters. *Journal of marketing*, 64(3):50–64.

Neslin, S. A. (2022). The omnichannel continuum: Integrating online and offline channels along the customer journey. *Journal of Retailing*, 98(1):111–132.

Nusrat, F. and Huang, Y. (2024). Feeling rewarded and entitled to be served: Understanding the influence of self-versus regular checkout on customer loyalty. *Journal of business research*, 170:114293.

Pillai, R., Sivathanu, B., and Dwivedi, Y. K. (2020). Shopping intention at AI-powered automated retail stores (AIPARS). *Journal of Retailing and Consumer Services*, 57:102207.

Santos Silva, J. M. C. and Tenreyro, S. (2006). The log of gravity. *The Review of Economics and Statistics*, 88(4):641–658.

Shriver, S. K. and Bollinger, B. (2022). Demand expansion and cannibalization effects from retail store entry: A structural analysis of multichannel demand. *Management Science*, 68(12):8829–8856.

Sinclair, W. and Stewart, H. (2026). Food accounted for 12.9 percent of U.S. households' expenditures in 2024. `https://www.ers.usda.gov/data-products/chart-gallery/58276`. Accessed: June 22, 2026.

Singh, V. P., Hansen, K. T., and Blattberg, R. C. (2006). Market entry and consumer behavior: An investigation of a Wal-Mart supercenter. *Marketing Science*, 25(5):457–476.

Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199.

Sun, Y., Wang, X., Hoegg, J., and Dahl, D. W. (2023). How consumers respond to embarrassing service encounters: A dehumanization perspective. *Journal of Marketing Research*, 60(4):646–664.

Ülkü, S., Hydock, C., and Cui, S. (2020). Making the wait worthwhile: Experiments on the effect of queueing on consumption. *Management Science*, 66(3):1149–1171.

Wang, K. and Goldfarb, A. (2017). Can offline stores drive online sales? *Journal of Marketing Research*, 54(5):706–719.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge, MA, 2nd edition.

Wooldridge, J. M. (2023). Simple approaches to nonlinear difference-in-differences with panel data. *The Econometrics Journal*, 26(3):C31–C66.

Xue, M., Hitt, L. M., and Chen, P.-Y. (2011). Determinants and outcomes of internet banking adoption. *Management science*, 57(2):291–307.

Zhang, J. Z., Chang, C.-W., and Neslin, S. A. (2022). How physical stores enhance customer value: The importance of product inspection depth. *Journal of Marketing*, 86(2):166–185.

Web Appendix

Contents

A Pre-Treatment Balance

A key concern in any difference-in-differences design is whether the treatment and control groups are comparable before treatment. Table A1 presents pre-treatment means for the treated and control groups across key shopping behavior variables, along with the difference and its statistical significance. We compute individual-level averages over each consumer's pre-adoption weeks, then compare group means. The groups are statistically indistinguishable on every shopping-behavior dimension; the one significant difference is the number of pre-treatment weeks observed (21.8 vs. 22.8, $p < 0.001$), which arises because the coarsened exact matching is on pre-adoption shopping behavior and city but not on the timing of the first CLS trip. The difference is small relative to the mean window length of roughly 22 weeks, and any resulting difference in calendar composition is absorbed by the week fixed effects in all specifications.

Variable	Treated	(SD)	Control	(SD)	Diff	p-value
Trips per week	2.11	(1.36)	2.12	(1.35)	-0.006	0.883
Spending per week (LCU)	29.37	(24.32)	29.53	(23.96)	-0.156	0.830
Items per week	6.29	(4.80)	6.34	(4.82)	-0.051	0.726
Pr(any trip in week)	0.72	(0.20)	0.73	(0.19)	-0.005	0.422
Spending per trip (LCU)	14.35	(7.77)	14.51	(7.63)	-0.154	0.505
Items per trip	3.05	(1.33)	3.09	(1.32)	-0.037	0.359
Pre-treatment weeks observed	21.84	(8.47)	22.78	(8.75)	-0.936	< 0.001
N (individuals)	2219		2219			

Table A1: Pre-Treatment Balance. The table compares mean pre-treatment shopping behavior between the treated group (consumers with more than one CLS trip) and the one-trip controls. Variables are computed as consumer-level averages over all pre-adoption weeks. p-values are from two-sample t-tests.

B Event Study Analysis

We estimate event study specifications that replace the post-treatment indicator in Equation 1 with a full set of relative-week dummies interacted with treatment intensity (reference week -1). Figure B1 plots the coefficients for weekly trips and spending. The pre-treatment coefficients are close to zero and statistically insignificant, indicating that treated and control consumers followed similar trajectories before adoption. After adoption the trip coefficient jumps and then declines over the following weeks. The spending response is more muted, consistent with trip splitting rather than consumption expansion.

Figure B2 shows that this decline is consistent with the gradual decline in the average usage of CLS over time—rather than a fading treatment effect. Figure B2 plots actual CLS trips per week among three terciles of consumer usage intensity. The High tercile holds steady at a little under one trip per week through week 11. The Low and Medium terciles fall off over the first month and then flatten out. The final point pools all weeks ≥ 12, where the sample is heavily skewed toward the earliest adopters as a third of treated consumers are no longer observed by that horizon. Heavy CLS users continue to visit CLS stores. The decline we observe in the event-study estimates (Figure B3) comes from the lighter two-thirds of treated consumers, whose usage settles at a low rate after an initial trial phase.

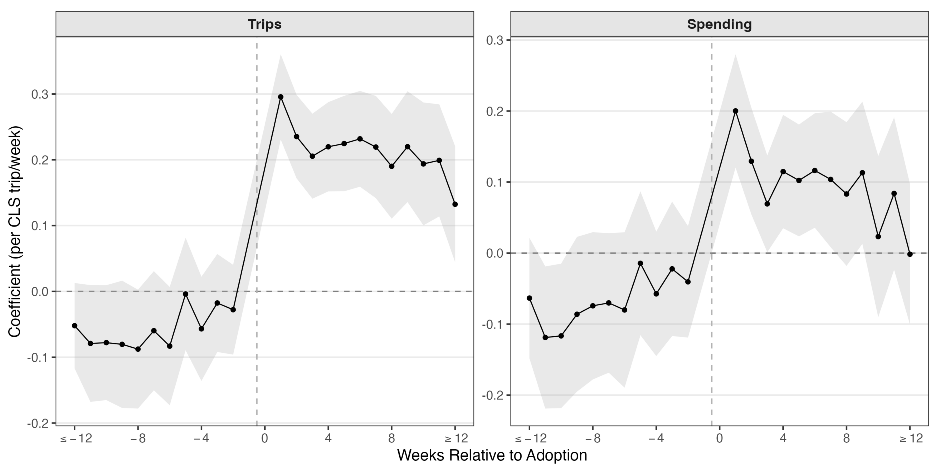

Figure B1: Event Study: Dynamic Treatment Effects of CLS Intensity on Weekly Outcomes. Each panel plots the coefficient on the event study of our main regression specification (Equation 1) on treatment intensity (average CLS trips per week). Relative week -1 is the reference period; the endpoint coefficients pool all weeks beyond ± 12. Shaded bands represent 95% confidence intervals. The vertical dashed line marks the adoption date.

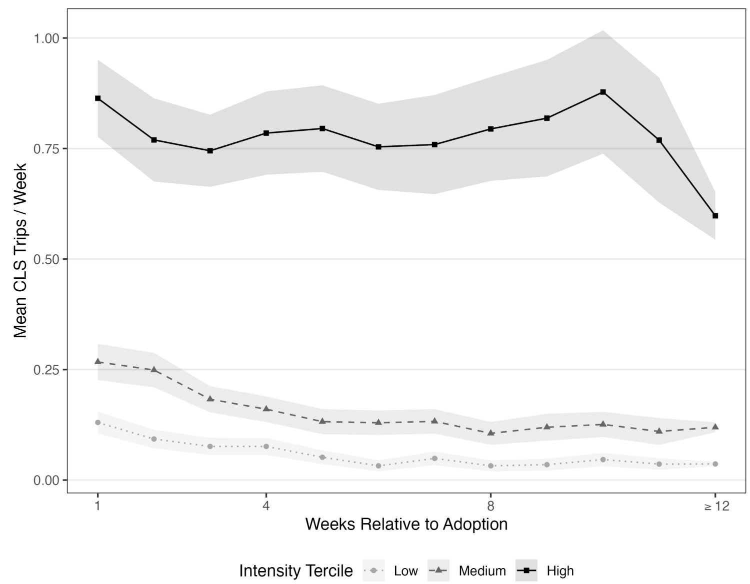

Figure B2: CLS Store Usage After Adoption by Intensity Tercile. Mean CLS trips per week among treated consumers, from week 1 onward; the adoption week is omitted. The final point pools all weeks ≥ 12.

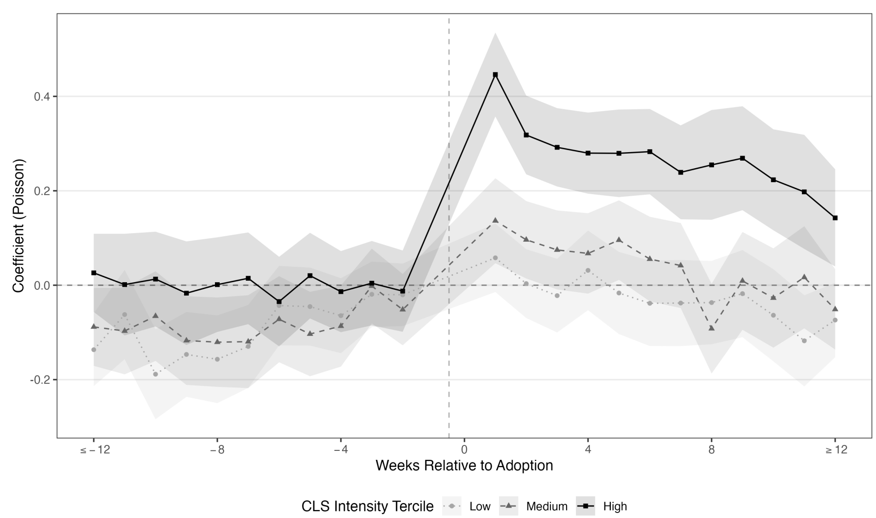

Figure B3: Event Study by CLS Usage Tercile. Each line plots Poisson event-study coefficients for one treated CLS-usage tercile (Low, Medium, or High); one-trip controls form the comparison group and are not plotted separately. Relative week -1 is omitted; endpoint coefficients pool weeks ≤ 12 and ≥ 12. Shaded bands show 95% confidence intervals, and the vertical dashed line marks adoption.

C Heterogeneity-Robust DiD Estimators

Adoption is staggered across consumers. Thus, we examine whether our conclusions hinge on the staggered, two-way fixed effects (TWFE) estimator. Table C1 re-estimates a binary treated/control version of the design with two cohort-robust staggered DiD estimators alongside the binary TWFE baseline. The first is the extended TWFE (ETWFE) estimator of Wooldridge (2023), developed for nonlinear difference-in-differences designs, including the Poisson case. The ETWFE estimator saturates the model with adoption-cohort-by-week treatment interactions and aggregates them into an average treatment effect on the treated (ATT) on the link scale. The second is the interaction-weighted estimator of Sun and Abraham (2021).

We rely on these estimators because both operate within the same Poisson fixed-effects framework as our main specification. The binary treatment is regular CLS users versus the one-trip controls. The adoption week itself (relative week 0) is excluded throughout, as in our main specifications. All three estimators recover a positive and statistically significant effect on weekly trips. The spending effects are positive but uniformly smaller, and statistically insignificant under Sun–Abraham.

Estimator	Trips ATT	Spending ATT
TWFE (binary, Poisson)	0.139*** (0.016)	0.072*** (0.019)
ETWFE (Poisson)	0.129*** (0.016)	0.057** (0.019)
Sun-Abraham (Poisson)	0.081*** (0.021)	0.024 (0.026)

Table C1: Heterogeneity-Robust Estimators: Average Treatment Effect on the Treated (ATT) on Weekly Trips and Spending. Binary treated (regular CLS users) versus one-trip controls; the adoption week (relative week 0) is excluded. All estimators are Poisson. Standard errors, clustered by consumer, in parentheses. $^{+}\,p < 0.1$; $^{*}\,p < 0.05$; $^{**}\,p < 0.01$; $^{***}\,p < 0.001$.

Each of these specifications requires analyzing binary treatments. Hence, these estimates lose the ability to analyze how treatment intensity influences the estimates, which we analyze in our main specification. Table C2 repeats the exercise contrasting high-intensity adopters (above-median CLS use) with the rest (low-intensity adopters and one-trip controls pooled). The estimated effects are uniformly larger than in the pooled comparison, consistent with the response concentrating among heavy users. Because the comparison group here includes low-intensity adopters, this estimand captures the effect of heavy use over and above light or no use, and we report it as secondary robustness only.

Estimator	Trips ATT	Spending ATT
TWFE (binary, Poisson)	0.201*** (0.020)	0.081** (0.025)
ETWFE (Poisson)	0.213*** (0.021)	0.086*** (0.024)
Sun-Abraham (Poisson)	0.198*** (0.031)	0.065+ (0.039)

Table C2: Secondary Robustness: Average Treatment Effect on the Treated (ATT) on Weekly Trips and Spending: High-Intensity Adopters versus the Rest. High-intensity treated (above-median CLS use) versus the rest (low-intensity adopters and one-trip controls); the adoption week (relative week 0) is excluded. All estimators are Poisson. Standard errors, clustered by consumer, in parentheses. $^{+}\,p < 0.1$; $^{*}\,p < 0.05$; $^{**}\,p < 0.01$; $^{***}\,p < 0.001$.

D Inter-Purchase Time

If CLS stores enable more frequent shopping, the time gap between consecutive trips should decline. We measure inter-purchase time (IPT) as the number of days between a consumer's successive store visits. Because weekly IPT is only defined for weeks with two or more trips—creating potential selection bias if treatment changes the probability of multi-trip weeks—we report two complementary approaches in Table D1. (i) A weekly panel regression (conditional on weeks with ≥ 2 trips), and (ii) a consumer-level comparison of average IPT in the pre- versus post-adoption period, which includes all consumers regardless of weekly trip frequency. Approach (ii) collapses time to two periods per consumer and estimates a consumer-fixed-effects panel with the same Post $\times$ Treated $\times$ intensity interaction, so its coefficients are directly comparable to the weekly panel in Column 1. It forgoes week fixed effects, but every consumer with trips in both periods enters the estimation, removing the multi-trip-week selection to which approach (i) is subject.

	Weekly IPT	Pre/Post Change
Post	-0.901***	0.821***
	(0.101)	(0.099)
Post:Treated	0.010	-0.100
	(0.130)	(0.146)
Post:Treated:CLSUsage	-1.915***	-1.603***
	(0.173)	(0.144)
Num.Obs.	117734	8876
R^2	0.201	0.711
RMSE	6.38	2.20
FE: Consumer	X	X
FE: Week	X	

Table D1: Estimates on Inter-Purchase Time (Days). Column 1: weekly-panel OLS, conditional on weeks with ≥ 2 trips. Column 2: two-period (pre/post) consumer-level panel. Standard errors, clustered by consumer, in parentheses. $^{+}\, p < 0.1$; $^{*}\, p < 0.05$; $^{**}\, p < 0.01$; $^{***}\, p < 0.001$.

Both approaches show that heavier CLS usage is associated with shorter time gaps between shopping trips. In the weekly panel, each additional CLS trip per week is associated with a decline in average inter-purchase time by about 1.9 days ($p < 0.001$). The consumer-level pre/post comparison estimates that each additional CLS trip per week is associated with a 1.6-day reduction in a consumer's average IPT ($p < 0.001$) across the entire post-period.

E Assortment and Store Choice

We examine how CLS usage is associated with changes in consumers' store-choice patterns. We estimate Equation 1 on three weekly consumer outcomes across all stores: (i) the average assortment of stores visited (number of unique products observed at each visited store in the data window); (ii) the number of distinct stores visited; and (iii) the share of trips to CLS stores.

	Avg Assortment	N Stores	Pct CLS
Post	-14.999**	0.033+	0.013***
	(5.770)	(0.018)	(0.001)
Post:Treated	14.056+	0.103***	0.032***
	(7.874)	(0.022)	(0.004)
Post:Treated:CLSUsage	-85.051***	0.141***	0.260***
	(11.804)	(0.030)	(0.011)
Num.Obs.	119532	119532	119532
R^2	0.480	0.250	0.404
RMSE	320.48	1.04	0.11
FE: Consumer	X	X	X
FE: Week	X	X	X

Table E1: CLS Usage, Store Assortment, and Store Choice. OLS regressions with individual and week fixed effects. Avg Assortment: mean number of unique products across stores visited that week. N Stores: distinct stores visited. Pct CLS: share of trips at CLS stores. Standard errors, clustered by consumer, in parentheses. $^{+}\, p < 0.1$; $^{*}\, p < 0.05$; $^{**}\, p < 0.01$; $^{***}\, p < 0.001$.

First, we observe that each additional CLS trip per week raises the share of weekly trips at CLS stores by 0.26 ($p < 0.001$), and consumers visit 0.14 more distinct stores per week ($p < 0.001$). This is consistent with CLS usage adding new trips and shopping occasions rather than fully displacing existing trips. Second, we see that the weekly average store assortment falls with CLS intensity (-85 unique products per CLS trip/week, $p < 0.001$).

At traditional stores only. If CLS stores open up new shopping occasions for CLS users, we further investigate whether this spills over into changes in shopping behavior at traditional stores. Table E2 re-estimates a similar set of outcomes on the sub-sample of non-CLS transactions only, along with trip-level basket measures at non-CLS stores (spending per trip, items per trip).

The treatment-intensity coefficient (Post $\times$ Treated $\times$ CLSUsage) is consistent with direct substitution: each additional CLS trip per week is associated with roughly 13% fewer traditional-store trips (Poisson coefficient -0.14, $p < 0.001$). Further, along the intensive margin at traditional stores, we observe 0.10 fewer items per traditional basket ($p < 0.05$) associated with each additional CLS trip per week.[6]

On the other hand, we see that the coefficients on store assortment (-4.2, $p > 0.1$) and the number of distinct traditional stores visited (-0.04, $p > 0.1$) are both small and insignificant with CLS usage. Thus, heavier CLS users do not appear to change which traditional stores they visit, but rather how often they shop and how much they buy on each trip. Importantly, this suggests no significant change in access to traditional stores associated with CLS usage.

Lastly, to further assess the reduction in the intensive margin at traditional stores, Table E3 replicates the volumetric analysis using purchases at traditional stores only. Consistent with fewer

[6] E.g., $\exp(-0.141) - 1 \approx -13\%$.

	Trad Trips	Trad Spent/Trip	Trad Items/Trip	Trad Avg Assortment	Trad N Stores
Post	0.035*	-0.263	-0.040	-8.928	0.014
	(0.017)	(0.162)	(0.029)	(5.819)	(0.017)
Post:Treated	0.035	0.041	0.068+	18.056*	0.051*
	(0.022)	(0.195)	(0.037)	(7.791)	(0.022)
Post:Treated:CLSUsage	-0.141***	-0.397	-0.100*	-4.222	-0.045
	(0.030)	(0.268)	(0.048)	(10.591)	(0.029)
Num.Obs.	165811	117792	117792	117792	117792
R^2	0.179	0.322	0.299	0.489	0.251
RMSE	1.94	10.64	1.91	323.19	1.03
FE: Consumer	X	X	X	X	X
FE: Week	X	X	X	X	X

Table E2: CLS Adoption and Traditional-Store Shopping Behavior. Regressions on non-CLS transactions only. Trad Trips is Poisson, including zero-trip weeks. The other columns are OLS regressions conditional on a traditional trip that week. Standard errors, clustered by consumer, in parentheses. $^+ p < 0.1$; $^* p < 0.05$; $^{**} p < 0.01$; $^{***} p < 0.001$.

traditional store visits, we find that each additional weekly CLS trip is associated with reduced total weight and volume purchased at traditional stores. We see directional evidence of per-item average weights and volumes reducing, but insignificantly so (point estimates of -2.6 g and -17 mL per item). These are substantially smaller than the overall per-item weight and volume reductions we find in the main analysis Table 8 (-6.6 g and -56 mL per item).

Overall, the results on store choice and traditional store shopping patterns are consistent with CLS adding a new shopping occasion that subtracts from existing traditional shopping in both frequency and basket size. This does not reshuffle where consumers do their full-basket trips as the assortment of traditional stores visited remains similar.

	Weight	Volume	Weight/Item	Volume/Item
Post	0.055*	0.080*	0.303	22.527
	(0.028)	(0.036)	(2.138)	(25.144)
Post:Treated	0.028	-0.009	-1.023	-22.167
	(0.034)	(0.040)	(2.815)	(24.338)
Post:Treated:CLSUsage	-0.186***	-0.197***	-2.583	-16.936
	(0.041)	(0.046)	(3.275)	(20.511)
Num.Obs.	165704	165811	86815	100753
R^2	0.323	0.325	0.174	0.106
RMSE	555.04	4372.48	131.55	1350.92
FE: Consumer	X	X	X	X
FE: Week	X	X	X	X

Table E3: Product Weight and Volume: Traditional Stores Only. Same construction as Table 8, restricted to purchases at traditional (non-CLS) stores: Columns 1–2 Poisson weekly totals; Columns 3–4 OLS per-item averages. Individual and week fixed effects. Standard errors, clustered by consumer, in parentheses. $^+ p < 0.1$; $^* p < 0.05$; $^{**} p < 0.01$; $^{***} p < 0.001$.

F Price Consistency Across Formats

One concern is that CLS stores may charge different prices for the same products, which could confound the treatment effects. We compare unit prices for the 1,131 products sold at both CLS and traditional stores. For each product, we average the unit price paid across all purchases at each format, pooling stores within a format, and compute the percentage difference between the CLS and traditional averages. Pricing is broadly consistent across formats (Table F1). The mean between-format gap is -0.69% (95% CI $[-1.46\%, 0.09\%]$), statistically indistinguishable from zero and far too small to account for the changes in trip frequency and basket size we document. This gap is also an order of magnitude smaller than the price variation consumers already face across traditional stores, where the same product's store-level average price varies by a median of 8.6% relative to the product's mean price. Price differences are therefore an unlikely driver of the behavioral changes we observe.

Statistic	Value
Products at both formats	1131
Mean price ratio (CLS/Traditional)	0.9931
Median price ratio	1.0232
Mean pct. difference [95% CI]	-0.69% [-1.46%, 0.09%]

Table F1: Price Consistency: CLS vs. Traditional Stores. Descriptive comparison of mean unit prices for products sold at both store formats. The price ratio is CLS price divided by traditional price; the percentage difference is reported with its 95% confidence interval across products.

G Within-Product Size Choice vs. Composition

Section 5.3 shows that CLS adopters buy items with lower per-unit volume and weight. We posit two channels that could drive this. Consumers may pick smaller packages of the same products they were already buying (*size choice*), or they may tilt their baskets toward products that simply come smaller (*composition*). Size choice may be related to changes in shopping occasions for CLS adopters, while composition could be related to both changes in shopping occasions or assortment. We conduct a few analyses.

G.1 Decomposition of Volume Decline

We first conduct a back-of-the-envelope decomposition of how much of CLS adopters' average decline in package volume is connected to size choice versus composition. The analysis uses product-volume data. Some products have multiple sizes, so we also identify "base" products that can come in different sizes, e.g. different sizes of a soda. For each product we convert their sizes, such as mL and cL, consistently into mL. We compare each consumer's pre-to-post change in average mL per item. This can be decomposed into two parts. We start with the pre-adoption period where we identify two aspects. One, we measure their pre-adoption product mix. Two, for each product in these shares, we can also calculate the average product volume. Subsequently, for the post-adoption comparison, we hold the consumer's product mix fixed and calculate the post-adoption average product volumes for those *same* products. This will explain some of the change in average product volume. The remainder will reflect a shift in product mix shares. For each consumer, we consider (base) products that they purchase in both periods.

We describe the decomposition results. We observe that treated consumers' average package volume falls by 6.1 mL per item. Across consumers, size choice accounts for about one-third (-1.7 mL) of this reduction. The remainder decomposes into composition changes for consumers (-4.4 mL). The one-trip control consumers have a different observed change in package volume; their average package volume increases by 11.4 mL.

G.2 Within-Consumer Size Choice

Following the back-of-the-envelope assessment, we conduct a formal regression analysis on a focused sub-sample that illustrates product size choice. Given the individual-basket level of the data, we can perform an analysis of product size choice at CLS stores versus at traditional stores. This analysis asks if the same consumer chooses a smaller product at CLS stores than they choose at traditional stores, even when CLS stores offer multiple options.

We consider base volumetric products that come in at least two distinct sizes, as long as each size has at least ten transactions in the data. This is 108 base products. Within this set, a consumer-base pair is a consumer who bought a given base at both formats post-adoption. There are 1,050 such pairs, spanning 871 consumers and 35 bases.

Each consumer-base pair thus contributes two observations, the consumer's mean package volume for that base at CLS versus at traditional stores. To conduct our analysis, we regress volume on a CLS indicator with individual and product-base fixed effects. Individual fixed effects absorb each shopper's overall size preference, product-base fixed effects absorb each product's typical size. The CLS coefficient then estimates the average within-person, within-product size choice gap between CLS and traditional formats. The sample is small because the design requires the identical consumer to buy the identical base product at both formats, but it seeks to isolate if size choices are format-specific.

	Multi-size bases	CLS-choice bases
CLS (vs. Traditional)	-141.88***	-133.08***
	(11.58)	(12.19)
Num.Obs.	2100	1548
R^2	0.768	0.805
RMSE	196.29	182.18
FE: Consumer	X	X
FE: base	X	X

Table G1: Within-Consumer Cross-Format Size Choice. OLS of average package volume (mL) on a CLS indicator, with consumer and base-product fixed effects; each consumer-base pair contributes its CLS mean and its traditional mean. Column 1: all multi-size pairs. Column 2: bases stocked in ≥ 2 sizes at CLS. Standard errors, clustered by consumer, in parentheses. $^{+}\, p < 0.1$; $^{*}\, p < 0.05$; $^{**}\, p < 0.01$; $^{***}\, p < 0.001$.

We find that the same consumer, when buying the same product, picks a package size that is on average 142 mL smaller at CLS than at a traditional store (Column 1). However, assortment is a concern. CLS may stock the smaller SKU, and consumers buy what is there. CLS stores are not the smallest format in the chain. The median CLS stocks roughly 308 unique products. While this is the 77th percentile of the traditional-store assortment distribution, the traditional stores that the sample consumers actually visit are larger with a mean of 797 unique products.

Column 2 addresses the assortment concern by restricting to 15 bases where CLS stores stock at least two distinct sizes, so the consumer has a real size choice at CLS stores. Surprisingly, we find the estimate is similar, at -133 mL (Column 2, $p < 0.001$). Given the small number of base products in the Column 2 analysis, we also provide a base-by-base comparison for transparency. Table G2 reports, for each of the 15 bases, the sizes CLS stocks, the number of consumer-base pairs, and the mean package volume the same consumers chose at each format. We manually present these bases into one group with small versus large options, versus a second group with only small options. Panel A shows bottled colas, waters, and juice are stocked in both single-serve (≤ 600 mL) and take-home (≥ 850 mL) variants. Panel B shows four bases that offer only single-serve variants: soft-drink metal cans (200/330 mL) and ready-to-drink coffees (200/400 mL).

The Panel B products are informative because it shows a small set of evidence against a uniform shelf-based explanation. If consumers simply reached for the smallest item on the shelf, cans may also show that because a 200 mL can is available at CLS. Instead, it appears that consumers buy the same 330 mL can at both formats. When a take-home size is available though, we observe that CLS adopters choose the smaller single-serving size more often at CLS stores. We read this as consistent with the occasion mechanism, with the caveat that the sample is small and confined to drinks.

Base product	Sizes at CLS (mL)	Pairs	CLS (mL)	Trad. (mL)	Diff. (mL)	p
Panel A: CLS stocks both single-serve and take-home sizes						
Orange juice, Brand A	300/1000	12	417	752	−335	0.014
Pepsi (PET)	500/850/1500	73	724	1011	−287	< 0.001
Still water, Brand B	600/1000/1500	65	1037	1300	−263	< 0.001
Still water, Brand C	500/1500	36	778	1026	−249	0.004
Pepsi Max Zero (PET)	500/850	55	643	847	−204	< 0.001
Still water, Brand D	500/1500	77	873	1051	−178	0.002
Coca-Cola (PET)	500/850/1500	119	778	945	−167	< 0.001
Mirinda Orange (PET)	500/1500	5	700	800	−100	0.374
Coca-Cola Cherry (PET)	500/850	33	550	650	−100	0.001
Coca-Cola Zero (PET)	500/850	69	624	705	−81	< 0.001
Sparkling water, Brand D	500/1500	38	1011	1014	−3	0.973
Panel B: CLS stocks only single-serve sizes						
Pepsi (can)	200/330	64	286	291	−5	0.498
Coffee with milk (ready-to-drink)	200/400	59	297	291	+6	0.601
Coca-Cola Zero (can)	200/330	61	299	267	+32	< 0.001
Flat white (ready-to-drink)	200/400	8	350	317	+33	0.170

Table G2: Per-Base Within-Consumer Size Comparison. For each of the 15 bases in Column 2 of Table G1: the distinct SKU volumes stocked at CLS ($\geq$ 10 transactions each), the number of consumer-base pairs (the same consumer purchasing the base at both formats post-adoption), and the mean package volume those consumers chose at each format. p-values are from two-sided paired t-tests across pairs within each base. Product names are translated from the local language; local-market brands are masked as Brand A–Brand D (Brand D covers the still and sparkling lines of one brand).

H Restricted Categories: Light and Heavy Segments

This section reports the per-category restricted regressions behind Figure 2 and Section 7.1.1. Table H1 estimates weekly category spending among all pre-adoption buyers of each restricted category; its coefficients are the spending estimates plotted for the four restricted categories in Figure 2. Tables H2 and H3 then split each category's buyers into Light and Heavy segments, below versus above the median of pre-treatment spending in that category.

	Beer	Wine	Strong Alcohols	Tobacco
Post	0.047	-0.180+	-0.196*	-0.018
	(0.051)	(0.094)	(0.093)	(0.048)
Post:Treated	-0.004	-0.087	-0.119	-0.031
	(0.064)	(0.111)	(0.112)	(0.062)
Post:Treated:CLSUsage	-0.139+	-0.135	-0.412***	-0.102
	(0.075)	(0.145)	(0.125)	(0.082)
Num.Obs.	122181	51779	66530	82335
R^2	0.292	0.200	0.209	0.362
RMSE	7.79	8.52	13.65	16.28
FE: Consumer	X	X	X	X
FE: Week	X	X	X	X

Table H1: Restricted Category Spending by Individual Category: All Buyers. Each column is a separate Poisson regression of weekly category spending, zero weeks retained, among consumers with positive pre-adoption spending in that category. These are the spending estimates shown for the restricted categories in Figure 2. Individual and week fixed effects. Standard errors, clustered by consumer, in parentheses. $^+ p < 0.1$; $^* p < 0.05$; $^{**} p < 0.01$; $^{***} p < 0.001$.

	Beer	Wine	Strong Alcohols	Tobacco
Post	0.468***	0.060	0.379*	0.489***
	(0.090)	(0.164)	(0.162)	(0.108)
Post:Treated	-0.096	0.037	0.013	-0.128
	(0.114)	(0.203)	(0.216)	(0.141)
Post:Treated:CLSUsage	-0.126	-0.490	-0.454+	-0.021
	(0.158)	(0.374)	(0.266)	(0.170)
Num.Obs.	59219	25779	32972	40569
R^2	0.143	0.117	0.149	0.222
RMSE	3.78	4.47	7.00	7.73
FE: Consumer	X	X	X	X
FE: Week	X	X	X	X

Table H2: Restricted Category Spending by Individual Category: Light Buyers. Each column is a separate Poisson regression for weekly category spending among Light buyers (below-median pre-treatment spending in that category). Individual and week fixed effects. Standard errors, clustered by consumer, in parentheses. $^+ p < 0.1$; $^* p < 0.05$; $^{**} p < 0.01$; $^{***} p < 0.001$.

	Beer	Wine	Strong Alcohols	Tobacco
Post	-0.043	-0.269*	-0.345**	-0.102*
	(0.059)	(0.111)	(0.107)	(0.052)
Post:Treated	0.013	-0.124	-0.168	-0.036
	(0.075)	(0.130)	(0.128)	(0.068)
Post:Treated:CLSUsage	-0.155+	-0.018	-0.375**	-0.099
	(0.085)	(0.143)	(0.134)	(0.092)
Num.Obs.	62962	26000	33558	41766
R^2	0.220	0.156	0.150	0.252
RMSE	10.21	11.16	17.88	21.48
FE: Consumer	X	X	X	X
FE: Week	X	X	X	X

Table H3: Restricted Category Spending by Individual Category: Heavy Buyers. Each column is a separate Poisson regression for weekly category spending among Heavy buyers (above-median pre-treatment spending in that category). Individual and week fixed effects. Standard errors, clustered by consumer, in parentheses. $^{+}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

I Trip-Frequency Terciles

This section reports a simpler cut of the shopping-style heterogeneity in Section 7.2. We split consumers into terciles of pre-treatment weekly trip frequency and estimate Equation 1 separately within each tercile, with weekly trips and spending as outcomes. Table I1 reports the estimates. The treatment-intensity coefficients decline monotonically with baseline frequency: 0.36 (Low), 0.27 (Medium), and 0.14 (High) for trips, with the same gradient for spending. Consumers who already shopped frequently before adoption show the smallest changes associated with CLS usage.

	Trips (Low)	Trips (Med)	Trips (High)	Spend (Low)	Spend (Med)	Spend (High)
Post	0.280***	0.070**	-0.036	0.264***	0.048	-0.017
	(0.035)	(0.026)	(0.026)	(0.053)	(0.031)	(0.036)
Post:Treated	0.037	0.016	0.022	-0.005	-0.016	-0.009
	(0.041)	(0.034)	(0.032)	(0.058)	(0.040)	(0.042)
Post:Treated:CLSUsage	0.359***	0.265***	0.143***	0.219***	0.165***	0.071+
	(0.026)	(0.040)	(0.028)	(0.033)	(0.048)	(0.038)
Num.Obs.	52515	57989	55307	52515	57989	55307
R^2	0.088	0.064	0.107	0.227	0.211	0.278
RMSE	1.21	1.69	2.65	26.86	31.79	48.12
FE: Consumer	X	X	X	X	X	X
FE: Week	X	X	X	X	X	X

Table I1: Treatment Heterogeneity by Pre-Treatment Trip-Frequency Tercile. Poisson regressions for weekly trips and spending, estimated separately by tercile of pre-treatment weekly trip frequency. Individual and week fixed effects. Standard errors, clustered by consumer, in parentheses. $^{+}\, p < 0.1$; $^{*}\, p < 0.05$; $^{**}\, p < 0.01$; $^{***}\, p < 0.001$.

J Hourly Shopping Profiles

This section supplements the hourly-profile clustering of Section 7.2. Table J1 reports the cluster-level estimates behind Figure 3 ($k = 4$).

	Trips (Morning)	Trips (Work)	Trips (Evening)	Trips (Night)	Spend (Morning)	Spend (Work)	Spend (Evening)	Spend (Night)
Post	0.034	0.012	0.035	0.107***	0.001	0.051	0.022	0.104*
	(0.047)	(0.028)	(0.036)	(0.032)	(0.059)	(0.035)	(0.047)	(0.045)
Post:Treated	-0.026	0.067+	0.033	-0.039	0.025	0.041	0.026	-0.111*
	(0.063)	(0.039)	(0.042)	(0.039)	(0.080)	(0.049)	(0.050)	(0.049)
Post:Treated:CLSUsage	0.357***	0.218***	0.219***	0.262***	0.219**	0.109*	0.132***	0.178***
	(0.069)	(0.037)	(0.039)	(0.053)	(0.074)	(0.049)	(0.038)	(0.049)
Num.Obs.	15674	51696	47602	49528	15674	51696	47602	49528
R^2	0.192	0.178	0.175	0.166	0.340	0.317	0.322	0.308
RMSE	1.95	2.02	1.92	1.94	32.48	35.05	37.05	39.92
FE: Consumer	X	X	X	X	X	X	X	X
FE: Week	X	X	X	X	X	X	X	X

Table J1: Treatment Heterogeneity by Hourly Shopping Profile (k-means, exploratory). Poisson regressions estimated separately by hourly-profile cluster ($k = 4$; see text). Individual and week fixed effects. Standard errors, clustered by consumer, in parentheses. $^{+}\,p < 0.1$; $^{*}\,p < 0.05$; $^{**}\,p < 0.01$; $^{***}\,p < 0.001$.

We also re-run the clustering with $k = 5$ on the same pre-adoption hourly trip-share profiles and re-estimate the treatment-intensity coefficient separately within each cluster. The extra cluster splits the broad work/midday group into two—an early-midday cluster peaking around noon and a late-midday cluster peaking in mid-afternoon—while the morning, evening, and night clusters are essentially unchanged.

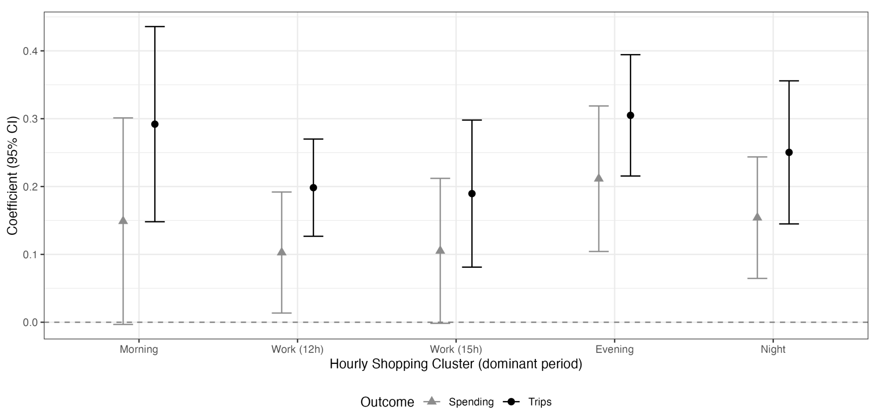

Figure J1: Treatment Heterogeneity by Hourly Shopping Profile, $k = 5$ (robustness). Consumers are clustered (k-means, $k = 5$) on their pre-adoption share of trips in each one-hour bin from 6am to 11pm; clusters are labeled by dominant time period, with the work/midday period splitting into a noon-peaking and a mid-afternoon-peaking cluster. Points show the treatment-intensity coefficient (Post $\times$ Treated $\times$ CLSUsage) for trips and spending, with 95% confidence intervals.

Both midday clusters—the noon- and mid-afternoon-peaking groups—exhibit the lowest trip-frequency coefficients (Figure J1), while the morning and evening clusters sit clearly above them. The midday trough is therefore robust to the finer partition.

K Robustness Checks

K.1 Instrumental variables: shift-share CLS access

A potential concern with our main specification is that treatment intensity is endogenous: consumers choose how much to use CLS, so unobserved changes in their circumstances could drive both usage and shopping outcomes. To address this, we construct a shift-share (Bartik-style) instrumental variable. The instrument measures how much CLS activity arrived in the neighborhoods where a consumer already shopped, using only where they shopped before their own first CLS trip (pre-adoption shares) and the chain-wide growth of CLS activity across districts. By using these aggregate measures, this separates the analysis from the consumer's own CLS usage.

The instrument exploits two sources of variation. The *share* component captures each consumer's predetermined geographic shopping pattern: using only that consumer's transactions from before their own first CLS trip, we calculate the fraction of their trips in each district. The *shift* component captures the time-varying rollout of CLS activity: for each district-week, we measure the fraction of trips occurring at CLS stores.

$$Z_{it} = \sum_{d} \underbrace{s_{id}^{\text{pre}}}_{\text{share}} \times \underbrace{\text{CLS.Pct}_{dt}}_{\text{shift}} \tag{K1}$$

where s_{id}^{pre} is consumer i's pre-adoption share of trips in district d, and CLS.Pct_{dt} is the percentage of trips in district d during week t at CLS stores. Because the shares are computed from pre-adoption behavior only, the consumer's own CLS trips never enter the instrument, and post-adoption changes in where they shop cannot feed back into it.

The shares are informative about CLS access only for consumers whose shopping geography is itself stable. If a consumer moves, changes routines, or splits time across cities, the districts visited before adoption are a poor guide to the access they actually faced afterward. We therefore estimate the IV on the subsample of consumers who are geographically *stable* throughout the study period, defined by two filters computed from their full trip histories: the same top shopping district in every month with at least three trips, and low drift between consecutive quarters of their trip history.[7] This retains 1,799 consumers. The filter is not mechanically related to treatment—stable and excluded consumers have nearly identical treated shares—and on this subsample the estimates are insensitive to how the share window is drawn: computing the shares from a fixed pre-launch calendar window instead of individual pre-adoption windows yields similar estimates on every outcome.

We estimate the IV in two steps. The first stage regresses actual CLS usage (Post $\times$ Treated $\times$ CLSUsage) on the instrument Z_{it} and the lower-order terms, with individual and week fixed effects. The second stage depends on the outcome. For the per-trip averages (OLS), we use 2SLS. For the weekly totals (Poisson) and the extensive margin (logit), the second stage is nonlinear, so we use a Control Function (CF) (Wooldridge, 2010, ch. 18). The CF keeps actual CLS usage in the second stage and adds the first-stage residual as a regressor. The coefficient on CLS usage is the IV estimate. The coefficient on the residual tests whether usage is endogenous. Standard errors on the CF usage coefficient come from a cluster bootstrap at the consumer level that re-estimates the first stage in each replication. Standard errors on the residual are analytic and clustered by consumer, which is valid for testing the null of exogeneity.

Table K1 reports the first stage and the IV estimates. The instrument predicts CLS usage with an F-statistic of 66 (Column 1). Instrumented CLS usage raises the probability of any weekly trip

[7] Drift is measured by the L_1 distance between the district shares of adjacent quarters of the trip history. We assume that stability means each distance to be at most 0.3. Based on L_1 distance, this can be interpreted as at most 15% of trip mass reallocating between quarters.

	First Stage	Trips	Spending	Items	Price/Item	Spending/Trip	Items/Trip	Pr(Any Trip)
Post:Treated:CLSAccess (IV)	2.881***							
	(0.354)							
Post	0.019***	0.055*	0.060+	0.058+	-0.030	-0.284	-0.055	0.100
	(0.003)	(0.025)	(0.032)	(0.030)	(0.055)	(0.242)	(0.045)	(0.062)
Post:Treated	0.195***	0.024	0.147+	0.107	0.326*	1.403*	0.186+	-0.448**
	(0.026)	(0.064)	(0.081)	(0.079)	(0.143)	(0.596)	(0.109)	(0.154)
Post:Treated:CLSUsage (IV)		0.177	-0.256	-0.106	-1.134***	-4.857***	-0.604*	1.557***
		(0.139)	(0.182)	(0.184)	(0.311)	(1.334)	(0.246)	(0.410)
First-stage residual		0.039	0.405*	0.307+				-0.830*
		(0.137)	(0.175)	(0.169)				(0.342)
Num.Obs.	69206	69206	69206	69206	53410	53410	53410	65323
R^2	0.609	0.168	0.338	0.266	0.219	0.352	0.331	0.146
Std.Errors	by consumer	see caption	see caption	see caption	by consumer	by consumer	by consumer	see caption
FE: Consumer	X	X	X	X	X	X	X	X
FE: Week	X	X	X	X	X	X	X	X
Estimator	OLS	CF (Poisson)	CF (Poisson)	CF (Poisson)	2SLS	2SLS	2SLS	CF (logit)
First-stage F	66.3	66.3	66.3	66.3	71.3	71.3	71.3	66.3

Table K1: Shift-Share IV: First Stage and IV Estimates. Column 1: first stage, OLS of the endogenous intensity interaction on the instrument and controls, with the F-statistic on the excluded instrument. Columns 2–8: IV estimates of the CLS usage effect. Weekly totals (Columns 2–4) and the extensive margin (Column 8) use the Control Function, with cluster-bootstrapped standard errors on the usage coefficient and clustered standard errors on the first-stage residual. Per-trip averages (Columns 5–7) use 2SLS with consumer-clustered standard errors. Individual and week fixed effects. Sample: the 1,799 geographically stable consumers defined in the text. $^+ p < 0.1$; $^* p < 0.05$; $^{**} p < 0.01$; $^{***} p < 0.001$.

(logit coefficient 1.56, $p < 0.001$). It reduces spending per trip by 4.86 LCU and price per item by 1.13 LCU (both $p < 0.001$), and items per trip by 0.60 ($p < 0.05$). The weekly totals are imprecise on this smaller sample. The trips, spending, and items estimates are insignificant. For trips and items, the confidence intervals include zero and include the treatment-intensity estimates from Table 5. We treat these margins as undetermined. The spending interval includes zero and the coefficient is directionally negative. This follows from the other coefficients, i.e., if trips do not clearly rise and each trip is smaller, weekly spending cannot rise by much.

The first-stage residual row reports the exogeneity test. For trips, the residual coefficient is 0.04 ($p = 0.77$). There is no evidence that the uninstrumented trips estimate is biased. For spending and items, the residual is positive (0.41, $p < 0.05$; 0.31, $p < 0.1$). This suggests that there may be unobservables that raise CLS usage and also raise spending, so the uninstrumented estimates may overstate the spending response. For the extensive margin, the residual is negative (-0.83, $p < 0.05$), so the uninstrumented estimate may understate the any-trip response.

The IV results are consistent with two findings from the main design using variation in access rather than choice. Consumers who use CLS because it is nearby shop in more weeks, consistent with the shorter gaps between trips in Section 6.2. They also buy smaller, cheaper baskets. These are the two margins at the center of the shopping-occasion mechanism (Section 6). The basket estimates are larger than their OLS counterparts. The IV complements the main design rather than replacing it. It identifies the response of geographically stable consumers whose usage is driven by proximity, on a restricted subsample. The weekly-total trip-frequency increase is directionally consistent, but imprecise, so the trip increase result relies on the within-adopter design of Section 4, which uses the full matched panel of CLS adopters.

K.2 Alternative control group

As an alternative to the within-adopter design, we replicate our results using a control group of consumers who never visited a CLS store, matched to treated consumers with the same coarsened exact matching procedure on pre-treatment shopping behavior used in the main sample ($N = 2{,}876$ per group). Because never-adopters have no CLS trip with which to date an event window, each control consumer is randomly assigned a synthetic adoption date drawn (with replacement) from the empirical distribution of adoption dates in the treated sample. The control group's pseudo-adoption timing therefore follows the calendar distribution of actual adoption. The same minimum-activity requirement around the (synthetic) adoption date is then applied to both groups. This design trades the within-adopter comparison's control over technology preferences for a comparison group with no CLS exposure at all: never-adopters need not share the treated group's awareness of, or willingness to try, the format, so we view this specification as a complementary robustness check rather than the primary design. Table K2 reports the weekly-outcome replication.

	Trips	Spending	Items	Price/Item	Spending/Trip	Items/Trip
Post	-0.002	-0.012	0.004	-0.043	-0.094	0.016
	(0.015)	(0.019)	(0.018)	(0.035)	(0.150)	(0.029)
Post:Treated	0.018	0.019	0.020	0.021	-0.080	-0.014
	(0.017)	(0.021)	(0.021)	(0.038)	(0.172)	(0.033)
Post:Treated:CLSUsage	0.269***	0.176***	0.210***	-0.376***	-1.673***	-0.197**
	(0.034)	(0.034)	(0.032)	(0.067)	(0.387)	(0.075)
Num.Obs.	216067	216067	216067	154119	154119	154119
R^2	0.181	0.330	0.272	0.181	0.331	0.308
RMSE	1.88	36.16	7.20	2.61	10.91	2.01
FE: Consumer	X	X	X	X	X	X
FE: Week	X	X	X	X	X	X

Table K2: Robustness: Weekly Outcomes (Alternative Sample). Never-adopter control group with synthetic adoption dates (see text). Columns 1–3: Poisson for weekly counts. Columns 4–6: OLS for per-trip averages. Individual and week fixed effects. Standard errors, clustered by consumer, in parentheses. $^{+}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

The treatment-intensity coefficients are close to their main-design counterparts in sign, significance, and magnitude: weekly trips 0.269 (main design: 0.238), spending 0.176 (0.148), and items 0.210 (0.195), with similar per-trip declines (spending per trip -1.67 versus -1.75 LCU). The main findings therefore do not depend on the choice between one-trip and never-adopter control groups.