BANNING UNSOLICITED STORE FLYERS: DOES HELPING THE ENVIRONMENT HURT RETAILING?

Jonne Y. Guyt[1]

Arjen van Lin

Kristopher O. Keller

September 2024

Jonne Y. Guyt[1] is an associate professor, Amsterdam Business School, University of Amsterdam, The Netherlands (e-mail: j.y.guyt@uva.nl).
Arjen van Lin is an associate professor, Tilburg School of Economics and Management, Tilburg University, The Netherlands (e-mail: avanlin@uvt.nl).
Kristopher O. Keller is an assistant professor, Kenan-Flagler Business School, University of North Carolina at Chapel Hill, USA (e-mail: kristopher_keller@kenan-flagler.unc.edu).

We thank Bart Bronnenberg, Els Gijsbrechts and participants at various conferences for helpful comments on an earlier version. We are indebted to AiMark/GfK for providing the data. We also thank the A Sustainable Future research platform and Marketing Science Institute (MSI) for financial support (MSI Research Grant # 4001614).

BANNING UNSOLICITED STORE FLYERS: DOES HELPING THE ENVIRONMENT HURT RETAILING?

ABSTRACT

Retailers often use store flyers to communicate the availability, price, and promotions of their products. Many households inspect store flyers for promotions, and the majority still prefer to receive store flyers in print. Yet, many are said to throw out store flyers unread, creating excessive waste that is environmentally damaging due to the excess use of paper, ink, and logistics. Legislation to reduce the distribution and waste associated with unsolicited store flyers has been proposed, but if and how much such a cut in distribution would affect households' grocery shopping behavior remains unclear. This paper investigates a recent policy change by seven Dutch municipalities that implemented a ban on unsolicited store flyers by moving from an opt-out to an opt-in policy at some point between 2018 and 2020. Using household scanner data, the authors assess changes in shopping behavior along nine comprehensive dimensions relevant to retailers, brand manufacturers, and policymakers, using a stacked, synthetic difference-in-differences approach. Although store flyer distribution decreased by 50% under the new policy, the drastic change did not substantially affect grocery shopping behavior. These findings are robust to different time windows and modeling approaches.

KEYWORDS: retailing, store flyer, natural experiment, advertising, responsible retailing, feature promotions.

INTRODUCTION

Store flyers, also known as store circulars or weekly ads, help retailers communicate about the availability, price, and promotions on the products in their assortment (Pieters, Wedel, and Zhang 2007). Even in the digital age, grocery retailers continue to invest in delivering physical store flyers to shoppers' homes, an investment that can reach $1 million for a single run, such that store flyers often account for the largest share of grocery retailers' marketing budgets (Kapner 2015b; Vericast 2020). In the United States, more than 80% of shoppers frequently inspect weekly flyers for promotions, and 60% prefer to receive store flyers in print format (Kapner 2015a; Vericast 2020). Such reliance is not unique to the United States either; store flyers account for more than half of the average retail marketing budget for grocers in several European countries (Gázquez-Abad and Martinez-Lopez 2016), where readership is also reported to be high. In the Netherlands, for example, 95% of all households that receive print store flyers read them, 86% use them, and 54% indicate they would miss them if the flyers stopped appearing in their mailboxes (GfK/NMO 2024).

But some environmental groups and policymakers doubt these claims, noting alternative evidence that many households throw out store flyers unread, creating unnecessary, excess uses of paper, ink, and logistics that have detrimental environmental implications (Retailtrends 2016). For example, in the Netherlands, a typical household receives approximately 1,300 store flyers annually, which could translate into 30 kilograms of paper waste per household (NOS 2021). In response to this concern, similar to several other national and regional governments (e.g., Australia, New Zealand, Quebec), the Netherlands has introduced legislation to reduce waste due to store flyers and other unsolicited advertising mail (Haas 2021; Woolf 2019). Others, such as France and Luxembourg, are in the process of introducing similar legislation (EUWID Pulp and Paper 2022; Near 2022).[1]

But, not all stakeholders agree on the environmental impact or that it comes without costs. For

[1] Most unsolicited advertising mail consists of store flyers, but current and proposed national policies address all its forms. In interviews with flyer distributors, we confirmed that most bulk mail involves store flyers, and retailers represent the majority of their clients. Hereinafter, reflecting our research focus, we use the term "store flyers".

example, in the Netherlands, several municipalities are hesitant to adopt the legislation, being worried about its net impact on society. Indeed, these plans have prompted strong opposition from flyer distributors, who fear that banning unsolicited store flyers will lead to negative downstream consequences in the distribution ecosystem, such as job losses. As importantly, such a ban may hurt retailers' business as a large percentage of their sales involve products promoted through the flyers (Gedenk, Neslin, and Ailawadi 2010). The effects of such a ban on household shopping behavior remains an open question, however, that neither trade press nor academic literature has answered. The effects on grocery shopping behavior will depend on the extent to which households rely on store flyers in practice. Industry reports suggest that households not only use store flyers to decide where to buy their groceries, but also what to buy (Grocery Dive 2018; Vericast 2020). For households using store flyers to guide their shopping, no longer receiving flyers can increase their search costs and may influence their grocery shopping behavior. Households often visit multiple retail chains in a given week (Gauri, Sudhir, and Talukdar 2008; Guyt and Gijsbrechts 2020) and may stop visiting some of these chains when they no longer receive their flyers and learn of their promotions. For instance, a household may purchase most of its groceries at its "primary" chain A but may purchase another product (e.g., cereal) at chain B in a particular week, having seen a price discount for that product in chain B's store flyer. Without receiving store flyers, the household may purchase all its groceries at its primary chain A that week. In addition, without flyers, the household may buy the private label product it always does instead of the brand advertised in chain B's store flyer or may not buy the product at all.

However, if environmental groups and policymakers are correct in that most households throw out store flyers unread, a ban's impact on shopping behavior is likely to be small. A common implementation of a ban on unsolicited store flyers is an opt-in approach in which only households that explicitly request flyers receive them. If households that use store flyers opt in to receive them and those who throw them out unread do not, waste can be reduced without any effects on shopping behavior. The policy change thus might act like an effective sorting mechanism.

Yet, we also might anticipate more nuanced outcomes because households make trade-offs between the benefits (e.g., learning about product availability, prices, and promotions, which can guide consumers' shopping) and costs of receiving store flyers. A default change could lead to higher costs (Johnson and Goldstein 2003), which might shift the balance of the cost-benefit trade-off—and thus households' shopping behavior. The costs of receiving flyers involve not just opportunity costs but also environmental costs and the potential or perceived costs of contradicting social norms (cf. Johnson and Goldstein 2003). Indeed, retailers and flyer distributors fear that households may stick to the default option of not receiving store flyers because the additional costs of are too high. That is, a default change may change the cost-benefit balance of receiving store flyers of a large group of households, leading to noticeable changes in shopping behavior. In this study, we aim to understand to what extent such concerns are warranted, i.e., whether a ban on unsolicited store flyers affects households' grocery shopping behavior.

To answer this pertinent research question, we consider a recent policy change by seven Dutch municipalities that implemented a ban on unsolicited store flyers, by switching from an opt-out to an opt-in policy for households, at some point between 2018 and 2020. That is, they moved from a setting where households receive store flyers unless they explicitly opt out to a setting where store flyers are not delivered unless households explicitly request them. Using household panel scanner data from GfK, we assess the change in shopping behaviors of about 900 households across the seven affected municipalities, using nine dimensions that are relevant to retailers, brand manufacturers, and policymakers: (1) total grocery expenditures, (2) total number of units purchased, (3) share of private labels, (4) share of purchases on promotion, (5) number of shopping trips, (6) number of different retail chains visited, (7) share of trips to traditional retailers versus hard discounters, (8) grocery expenditure at the primary chain, and (9) number of shopping trips to the primary chain.

As an identification strategy, we use the recently developed, stacked, synthetic difference-in-differences (SynthDID) approach (Arkhangelsky et al. 2021; Berman and Israeli 2022; Li et al. 2022) that enables us to compare households' shopping behavior before and after the policy change, relative to a set of control households in markets without the ban (i.e., still subject to the opt-out policy). The stacked

SynthDID approach offers several advantages in our setting. First, it creates a synthetic control group, matched on pre-treatment purchase behavior, which enhances the compatibility and face validity of the control households vis-à-vis the treatment households. Second, it comes with a desirable bias-variance tradeoff compared to alternative methods such as DID or synthetic control. Third, it reduces the risk of overfitting that can occur with synthetic control. Fourth, it accounts for the staggered implementation of the ban across municipalities, which renders the standard DID approach biased (Goodman-Bacon 2021).

The findings show that instituting an opt-in policy reduces the distribution of store flyers dramatically: Under the previous opt-out policy, close to 75% of all households received store flyers, whereas that level fell to 37.5% under the opt-in policy. Contrary to concerns by retailers and flyer distributors, the stacked SynthDID findings show that this drastic change does not substantially affect any of the nine dimensions of grocery shopping behavior of households in the treated areas. That is, the intent-to-treat (ITT) estimates are statistically not distinguishable from zero and are economically small. These findings remain remarkably robust to different time windows and modeling approaches. Results from a subgroup analysis, in which we estimate the local average treatment effect (LATE) for households that did not opt in, i.e., that stopped receiving store flyers, similarly show no economically meaningful change in behavior. Specifically, for eight of our nine dimensions of grocery shopping, we find small and insignificant effects, while for the share of purchases on promotions, we find a significant but economically small effect of less than one percentage point decrease. We also provide policy-relevant insights as to whether some households are differentially affected from others by this legislative change. Our results provide limited evidence of this; only older and moderately price-sensitive households appear to be impacted as they modestly reduce their share of purchases on promotions. We explore three explanations for this lack of meaningful effects: Retailers changing their marketing communication in treated municipalities, households switching to digital rather than print flyers, or an overstated dependence of some households on store flyers. We find no evidence for the first two explanations. As such, our findings suggest that the 50% of households that no longer receive flyers after the ban did not rely on them and, hence, were largely unresponsive to these marketing efforts.

Thus, this study contributes to academic literature, practice, and policy. Store flyers and feature promotions have been analyzed extensively in marketing literature (Ailawadi and Gupta 2014), primarily in efforts to understand the implications of featuring certain products in store flyers (e.g., Guyt and Gijsbrechts 2018; Van Lin and Gijsbrechts 2016) or establish the role of store flyers in the conversion funnel (Seiler and Yao 2017). No study considers a situation in which households do not receive any unsolicited flyers at all. Our unique research setting makes it possible to ascertain reactions to banning unsolicited store flyers in a quasi-experimental way. Importantly, we study a ban on *unsolicited* store flyers through a change in the default option, not an outright ban, which forbids distribution to all households. As such, we evaluate to what extent retailers are hurt by an increasingly common policy where interested households could still opt in, rather than evaluate the value of store flyers themselves.

Managerially, our findings provide relevant insights for both retailers and brand manufacturers. First, retailers that worry about a ban on unsolicited story flyers benefit from understanding its business impact. This is particularly relevant for traditional retailers, for which store flyers often are the main vehicle to communicate deals and promotions. Second, national brand manufacturers also rely heavily on feature promotions and may also lose from a ban. Yet our results show that neither retailers nor manufacturers are meaningfully affected by a ban on unsolicited store flyers. These insights help retailers and manufacturers facing such bans understand their impact with greater confidence.

Finally, our results are relevant for policymakers. A ban on unsolicited store flyers could reduce substantial amounts of paper used as well as the energy and water used to produce and transport flyers. Yet responsible policymakers also must acknowledge potentially unintended consequences of the legislation they propose, including potential adverse downstream consequences for flyer distributors, less than expected materialization of environmental benefits, and the risk of adverse implications of an opt-in default for grocers, which still heavily rely on store flyers. Critically, our findings indicate that at least concerns regarding changes in shopping behavior appear unfounded.

RELEVANT LITERATURE

Store Flyers and Feature Advertising

Our study relates and contributes to literature on store flyers and the promotions featured in these flyers, which also are known as feature ads. Store flyers are distributed regularly, often weekly or bi-weekly, to households, and communicate the availability of products (i.e., a retailer's assortment) and information on current prices (Pieters et al. 2007). As such, consumers may use store flyers for various reasons, including deciding what and how much to buy and which store to visit (Gauri et al. 2017). Store flyers are a form of advertising mail, and in terms of their effect on consumers' decision process (the "mechanism"), store flyers and the information provided therein can guide consumers' shopping behavior by creating awareness for promoted products and lowering search costs (similar to the effects of advertising; Honka, Hortaçsu, and Wildenbeest 2019). That is, the information provided in the store flyer makes it easier to identify the best deals, and to determine what to buy and which stores to visit, such that they may, for instance, buy a certain national brand vs. a private label product or visit an additional retail chain to benefit from its current promotions.

Academic research has shown that store flyers drive store traffic and sales (e.g., Gabel, Molitor, and Spann 2024; Guyt and Gijsbrechts 2020). For instance, Gabel, Molitor, and Spann (2024) show how households shift their purchases from retailers that use store flyers to retailers that do not use them during a three-week suspension during COVID-19. In further studying its value, many studies consider the composition of the store flyer (e.g., size, allocation of space to particular categories or brand types, average depth of the discount; Gijsbrechts, Campo, and Goossens 2003; Gázquez-Abad and Martinez-Lopez 2016). For example, Gijsbrechts, Campo, and Goossens (2003) show that store flyers featuring deeper discounts are more effective for driving sales and traffic. Another subset of studies addresses the effect of individual feature promotions in the store flyers, which can increase both product sales and category purchases (e.g., Fox and Hoch 2005; Guyt and Gijsbrechts 2018). Their impact depends on specific factors, such as discount depth (van Lin and Gijsbrechts 2016) and the type of retailer (traditional versus hard discounter; Widdecke et al. 2023) or product (private labels vs. national brands; Lemon and Nowlis 2002). That is, some feature promotions may be more relevant for certain products or retailers.

Providing a more nuanced picture, Srinivasan and Bodapati (2006) show that feature advertising drives the store choice decisions of only certain consumers. Indeed, several studies indicate that not all consumers who receive store flyers use them to decide on their shopping (Gázquez-Abad et al. 2014; van Lin and Gijsbrechts 2016). Moreover, using path-tracking data, Seiler and Yao (2017) find that feature advertising prompts existing consumers to buy more of the featured brand rather than encouraging more consumers to visit the product category.

However, among these various investigations of store flyers, their feature promotions, and their effects on consumers' grocery shopping behavior, we find no study that addresses the effects of banning unsolicited store flyers. Summarizing the existing literature, two abstract findings can guide our intuition for the implications of a ban on unsolicited store flyers. First, store flyers can direct households' shopping in terms of the choice of store and what and how much to buy, i.e., basket composition within a store. As such, it is important to cover a wide set of grocery shopping dimensions to understand the impact of a default change. Second, there appears to be considerable heterogeneity in the usage of the store flyer. In our setting, we also expect that not all households value store flyers equally and differ in whether they use them to guide their shopping. Some households may not use them at all because they see little benefit in the provided information, whereas others may use flyers more intensively. In the next section we detail how a change in the default (i.e., from opt-out to opt-in) may translate into material changes to the shopping behavior of households.

The Effect of a Default Change

If households appreciate and want to use store flyers, a ban on unsolicited advertising would require them to opt in to continue receiving them. The policy change thus might act like a sorting mechanism, with little effect on behavior, because only households that did not use store flyers (but received them by default) stop receiving them, whereas households that derive benefits from them opt in and continue to receive them. Yet we also might anticipate more nuanced outcomes, because households make trade-offs between the benefits and costs of receiving store flyers. A default change could lead to higher costs (Johnson and Goldstein 2003), which might shift the balance of the cost-benefit trade-off—and thus households' shopping behavior. We discuss the effects below, after which we discuss the size of any potential imperfect sorting.

As explained above, the benefits of receiving print store flyers for households include learning about product availability, prices, and promotions, which can help consumers decide what and how much to buy and which store to visit. The costs of receiving flyers involve not just opportunity costs but also environmental costs and the potential or perceived costs of contradicting social norms (cf. Johnson and Goldstein 2003). For example, with regard to the opportunity costs, households might resist the cognitive effort required to choose whether to (still) receive store flyers,[2] in line with descriptions of people's status quo bias (Fernandez and Rodrik 1991; Samuelson and Zeckhauser 1988). Such effort had not been required previously, because households received store flyers by default. Once those flyers stop arriving, households must undertake some non-trivial cognitive effort to evaluate the benefits, such as the potential to learn about the retailer's assortment or opportunities to save by reviewing the printed store flyers. If the cognitive efforts seem likely to outweigh the benefits, they might accept the new default option of not receiving store flyers. Furthermore, to obtain a mailbox sticker (e.g., filling in a form to receive it, learning where to obtain it, seeking it out at a local office) and preparing their mailbox, households must engage in logistical efforts (Stantec 2024). According to the Dutch Directorate General for Public Works and Water Management, some households express aesthetic concerns about stickers on their mailbox too (Stantec 2024). Prior to the ban, these costs only applied to households that actively sought to opt out; after the ban, households would incur the costs of receiving store flyers. If the costs outweigh the perceived benefits, households likely accept the new default and no longer receive store flyers, even though they might derive benefits from them.

[2] A report commissioned by the Dutch Directorate General for Public Works and Water Management (Stantec 2024) indicates that about 10% of all households do not know which policy (opt-in or opt-out) is in effect in their municipality, giving credence to the claim that households might avoid investing effort into understanding whether they should deviate from the default option.

Default changes also offer institutional signals of societal norms (Tankard and Paluck 2016) and an implicit recommendation about whether receiving store flyers is (still) societally preferable. When receiving store flyers is the default (i.e., opt-out regime), households have little reason to pay attention to the negative (environmental) effects of receiving store flyers. When not receiving store flyers becomes the default (i.e., opt-in regime), households likely reevaluate and recognize the potential negative effects, guided by environmental regulations (Mettler and Soss 2004; Schultz 1999). According to queue theory, changing the default option affects the order and type of arguments considered, prompting earlier retrieval of more positive aspects of the default and negative features of the non-default option (Dinner et al. 2011). In turn, we predict that the benefits provided by the store flyer get considered earlier in an opt-out regime, but environmental gains (by not receiving store flyers) are considered more in an opt-in regime. If social perceptions then suggest that continuing to receive print store flyers (which eventually end up in the trash) is not socially acceptable, they might lead to social judgments. If the perceived costs of receiving store flyers increase due to such reevaluations of their environmental impact and social frictions, more households might accept the new default, even if they find the store flyer useful.

We predict that these cost shifts, due to the default change, affect many households and their decision to accept the new default or opt in to continue receiving flyers. But for some households, such as those at extreme ends of the cost–benefit trade-off, such changes appear less likely. These households are either very positive or very negative about receiving store flyers, so they either opt in after the regulatory change or have opted out previously. But because the cost increase can overwhelm the benefits of store flyers for shoppers who fall somewhere between these extremes they may not choose to opt in (i.e., no longer receive store flyers), so the net balance of the trade-off could shift for a non-trivial group of households.

Empirically, there is evidence of default changes resulting in large shifts in other settings, such as organ donation (from 82 to 42%; Johnson and Goldstein 2003). This, combined with survey statistics suggesting only roughly 50% would opt-in out of the households that currently receive store flyers (Markteffect 2022), suggests many households may experience a shift in their net balance and stop receiving store flyers once an opt-in policy is in place. A priori, it is unclear whether these households predominantly consist of households that obtain few benefits from the flyer, as policymakers suggest, or households that do obtain substantial benefits, as suggested by the industry. Importantly, according to supporters of the opt-in policy, the opt-out policy leads to imperfect sorting as there are many households that are receiving but not using flyers, whereas opponents typically argue that an opt-in policy leads to imperfect sorting as many households derive benefits from flyers but would stop receiving them.

To summarize, the degree to which shopping behaviors change due to the new default depends on the share of households that see a shift in their net balance and the degree to which they derive benefits from store flyers that drive their shopping behavior, which are empirical questions.

SETTING AND DATA

***Setting and Institutional Details*.**

The setting of our study is the grocery market in the Netherlands, where many retailers rely on store flyers as one of their main communication vehicles.[3] Between 2018 and 2020, seven municipalities introduced a ban on unsolicited store flyers. An opt-out approach has been in place since the early 1990s and about 20%–30% of Dutch households had opted out of receiving store flyers. But environmental groups and local politicians questioned its effectiveness. By 2016, Amsterdam was the first municipality to approve a ban, such that its opt-in policy came into effect in January 2018. Flyer distributors questioned the legality of the policy and filed lawsuits, which delayed its approval and introduction in other municipalities. The courts ultimately ruled in favor of the municipalities, leading to the new policy implementation in Haarlem, Leiden, Tilburg, and Utrecht in January 2020; The Hague in July 2020; and Rotterdam in November 2020. However, several other municipalities have not banned unsolicited store flyers thus far and retained the opt-out policy, as depicted in Figure 1.

[3] All but one of the national retail chains use store flyers, and, across the other chains, there was no variation in distribution across municipalities before the ban. A separate analysis that assesses shopping behavior at this chain vs. others shows that the new policy does not affect this chain differently.

FIGURE 1: MUNICIPALITIES THAT BANNED UNSOLICITED STORE FLYERS (GREY) OR NOT (WHITE) IN THE NETHERLANDS UNTIL THE END OF 2021

The most important driver in favor of adopting the opt-in policy, according to municipalities, is whether it could lead to a meaningful reduction in paper and environmental waste related to the printing and distribution of store flyers (Stantec 2024). Other drivers mentioned are a (potentially sizable) reduction in garbage collection costs and stimulating environmental awareness among its citizens. Municipalities that have not (yet) introduced a ban question the effectiveness of the policy in reducing paper and environmental waste and may expect other measures to be effective in that regard too. Additionally, municipalities differently evaluate how feasible implementation and enforcement of the policy is. Another factor is the degree to which municipalities evaluate the risks in terms of local business (and job losses). Lastly, municipalities have been cited as ‘not wanting the hassle’ related to implementation in the form of legal issues (Stantec 2024). Indeed, as indicated above, the first municipality to introduce a ban (Amsterdam) dealt with numerous lawsuits, which may be too costly for smaller municipalities to bear.[4]

[4] Rotterdam, Utrecht, and the Hague also dealt with lawsuits surrounding the implementation (and its details). For an overview of all lawsuits, see Stantec (2024, appendix B).

Mailbox stickers indicate households' opt-out/in preferences, i.e., the treatment is at the household level. Before the ban was introduced, households could put a "NO" sticker on their mailbox, indicating they did not want to receive any unsolicited store flyers. After its introduction, households could put a "YES" sticker on their mailbox to indicate that they still want to receive such mail. The stickers are available online or at the local city hall, free of charge. Any previously used "NO" stickers remain valid but are equivalent to not using a sticker. Table 1 clearly details whether households receive store flyers (✓) or not (✗) according to these two approaches.[5]

TABLE 1: RECEIVING STORE FLYERS (✓) OR NOT (✗) PRE AND POST BAN

		Pre-ban	Post-ban
Mailbox sticker usage	**Without sticker**	✓	✗
	"NO" sticker	✗	✗
	"YES" sticker	n/a	✓

Data

We use GfK household scanner panel data for the period 2017–2021. These data contain purchase information of more than 10,000 households in any given year, representing a stratified national sample; for each trip made, they reveal the specific retail chain visited and exact purchase records (i.e., items and quantities bought, prices paid, whether the item is a private label product, and whether the item was purchased on promotion). We use these data to compute households' weekly shopping behavior along nine dimensions: (1) total grocery expenditure, (2) total number of units purchased, (3) share of private labels, (4) share of purchases on promotion, (5) number of shopping trips, (6) number of different retail chains visited, (7) share of trips to traditional retailers versus hard discounters, (8) grocery expenditure at the primary chain, and (9) number of shopping trips to the primary chain. The share-based measures and the number of different retail chains visited are computed conditional on a trip observed in a given week. Next to their purchases, panelists also provide information about their demographics and home location (four-digit zip code and municipality). About 12% of the households live in one of the seven treated

[5] In addition to indicating their preferences regarding unsolicited advertising mail, households can indicate their choice to receive community newspapers separately; the new policy does not apply to community newspapers. We provide more information on the implementation and different stickers in Web Appendix A.

municipalities. To understand how the ban affects their receipt of store flyers, we also need data about their opt-out/in choices (i.e., use of a mailbox sticker). We obtain these data from GfK and NOM, a Dutch media research organization, which survey a sample of panel households annually. In the first week of 2021, just after the latest introduction of the ban, we ran a survey of all treated households to supplement the data from GfK/NOM, asking households if they used a mailbox sticker and when they started using it. We obtained information about the decision to (not) opt in for about 60% of the treated households.[6]

Model-Free Evidence

Before presenting our model, we provide some descriptive statistics and model-free evidence regarding the distribution of store flyers and household shopping behavior. First, in terms of circulation, we observe that 37.5% of the households in municipalities with a ban opted in to receive the flyers. Prior to the ban, we observe an opt-out of only 25.7%, implying that 74.3% of the households received flyers before the ban was introduced. Therefore, the number of flyers distributed declined by about 49.5% (from 74.3% to 37.5%). Second, when we compare households' shopping behavior 52 weeks before versus after the ban, across all households living in the treatment municipalities, we find only small differences in the nine grocery shopping behaviors, even though circulation dropped by almost half (see Table 2, Panel A). The largest difference we observe is for the average number of shopping trips to the primary retail chain, which decreases from 1.49 to 1.39 (-6.7%). Yet, other developments could have affected household shopping behavior in the study period too. Indeed, a comparison of shopping behavior of households from municipalities where a ban was not (yet) present across the same period (in Table 2, Panel B) shows a similar decrease. Despite retailers' concerns that reducing their flyers' distribution would hurt their business, this evidence suggests that the ban had only marginal effects on household shopping behavior. Still, the typical caveats with before and after comparisons apply, and we address these next.

[6] A comparison of the demographics of households for which this information is available vs. not shows negligible differences between the two groups in terms of age, household size, and social class. However, the information is more likely to be available for households in higher income brackets. In our analysis, we use a random forest algorithm to impute the sticker usage of other households using households' income class as input (amongst other variables). As alternative, we condition on households for which we have the information in our analysis.

TABLE 2: DESCRIPTIVE STATISTICS HOUSEHOLD SHOPPING BEHAVIOR

A: HOUSEHOLDS IN TREATED MUNICIPALITIES

	Pre-Ban		Post-Ban	
	Mean	SD	Mean	SD
Total grocery expenditure (in euros)	55.08	40.58	58.08	42.70
Total number of units purchased	45.02	35.66	45.68	35.75
Share of private labels (%)	53.71	22.92	53.47	22.90
Share of purchases on promotion (%)	21.58	18.50	21.17	17.94
Number of shopping trips	2.59	1.88	2.46	1.78
Number of different retail chains visited	1.96	1.04	1.91	1.02
Trip share HDs vs. traditional retailers (%)	17.55	27.13	17.16	27.40
Grocery expenditure at primary chain (in euros)	37.44	34.98	38.18	37.28
Number of shopping trips to primary chain	1.49	1.19	1.39	1.19

B: HOUSEHOLDS IN CONTROL MUNICIPALITIES

	Pre-Ban		Post-Ban	
	Mean	SD	Mean	SD
Total grocery expenditure (in euros)	60.30	41.54	62.17	43.06
Total number of units purchased	50.46	37.17	50.44	37.05
Share of private labels (%)	55.98	23.57	56.06	23.44
Share of purchases on promotion (%)	20.24	17.73	20.28	17.52
Number of shopping trips	2.64	1.85	2.49	1.79
Number of different retail chains visited	2.06	1.05	2.01	1.04
Trip share HDs vs. traditional retailers (%)	23.77	29.49	23.66	30.00
Grocery expenditure at primary chain (in euros)	39.88	35.17	39.82	36.85
Number of shopping trips to primary chain	1.46	1.14	1.35	1.11

Notes: Pre-ban and post-ban refer to 52 weeks before and after the ban, respectively. Number of treated (control) households: 892 (32,303) for variables unconditional on a trip observed in a week, and 361 (13,474) for variables conditional on a trip observed in a week. For a description of the sample selection process see the section "Estimation Sample (Treated and Control Households)". Share of private labels is the percentage of private label units, out of total units purchased in a given week. Share of purchases on promotion is computed in a similar way. HD = hard discounter. Aldi and Lidl are coded as HDs, all other retailers as traditional retailers. The primary chain is defined as the chain to which the household allocated most of its expenditure prior to the introduction of the ban.

METHOD

We analyze the impact of introducing a ban on unsolicited store flyers by comparing changes in households' shopping behavior using households from municipalities where a ban was introduced to households from municipalities where such a ban was not (yet) present. In terms of the comparative approach, two characteristics of our setting require special care. First, as Figure 1 indicates, the adoption of the ban has been limited and concentrated in larger cities in the western part of the Netherlands. The average household from the control municipalities thus may not only differ in location but may also differ in other dimensions relevant to grocery shopping behavior compared to households in the treatment municipalities. Because our approach relies on a before-after-with-control household framework, we need an approach that ensures a good comparison group. Second, the ban was introduced in a staggered manner, so we need a model that can accommodate staggered treatments (Goodman-Bacon 2021). Before discussing our comparative approach in detail, we discuss how we select an identification strategy suitable for our study setting.

Identification Strategy

Commonplace methods to analyze such settings are the difference-in-differences (DID) method, the synthetic control (SC) method, and the recently proposed synthetic difference-in-differences (SynthDID) method. While these models all rely on a parallel trends assumption,[7] these models differ in how control units and time periods are selected, reflected in different model restrictions. We summarize these differences in Table 3. The regular DID places the strongest restrictions by weighting all controls equally. The SC method relaxes this restriction and creates a weighted combination of controls with weights chosen such that the average pre-treatment outcome of the controls is approximately equal to the average pre-treatment outcome of the treated units (with the weights restricted to be positive and to sum to one and without allowing for a baseline difference between treated and control units, i.e., a "zero intercept" restriction; Li and Shankar 2023). Importantly, in our case, the treatment is not randomly allocated to municipalities in the Netherlands and tends to focus on the larger ones. Households living in control municipalities may be different from households living in treatment municipalities and exhibit different shopping behavior, on average. Given these differences, equally weighting all households (as in DID) is theoretically unappealing, and a method like SC that flexibly adjusts the weight of each control household would be preferred. Indeed, by "dropping some controls that are (often weakly) negatively correlated or uncorrelated with the treated unit" (Li 2020, p. 2068), SC can reduce the estimation variance compared to DID. At the same time, as Arkhangelsky et al. (2021) note, SC weights tend to be sparse

[7] While DID relies on the assumption that a treated unit's outcome would have been parallel to the average control unit's outcomes in the absence of the treatment, both SC and SynthDID rely on the assumption that a treated unit's outcome would have been parallel to a weighted average control unit's outcome in the absence of the treatment. An untestable part of the identifying assumptions is that the correlation structure between treatment and (potentially weighted) control units continues into the post-treatment period.

such that a few controls may be very influential (see their Figure 1, p. 4095). Yet, these few controls may be matched on noise (i.e., the SC method may overfit in the case of many potential controls; see Abadie 2021 and Li and Sonnier 2023), challenging the validity of the counterfactual outcome, and thereby, the validity of the estimated treatment effect. The synthetic difference-in-differences (SynthDID) method proposed by Arkhangelsky et al. (2021) is attractive because it introduces a regularization parameter to increase the dispersion of the weights (such that none of the controls have a particularly strong influence), which is particularly attractive in the case of many potential controls (as is the case in our setting). It also allows for baseline differences between treated and control units (relaxing the "zero intercept" restriction of the SC method), which, together with the regularization parameter, enables a more balanced weighting with less sparse weights (Arkhangelsky et al. 2021, p. 4094). Next to that, the SynthDID method also introduces time weights next to unit weights that can further "remove bias and improve precision by eliminating [pre-treament] time periods that are very different from post-treatment periods" (Arkhangelsky et al. 2021, p. 4090), relaxing the equal weights or zero weights restriction of DID and SC, respectively.[8] As Arkhangelsky et al. (2019, 2021) show, the SynthDID method comes with a good bias-variance tradeoff and the additional parameters come at limited costs of additional variance.[9] Nevertheless, we note that estimates stemming from more restrictive approaches should be more precise if their respective assumptions hold.

Last, we note that in the case of a staggered adoption we rely on an additional assumption, namely that of adoption exogeneity, which deals with the order of the treatment across cohorts. More specifically, the effect of each treatment should be mean independent of the cohort treatment sequence.

[8] While our setting provides a theoretical rationale for the usage of unit weights, an equally strong motivation for the time weights is absent. In our main model, and based on the empirical findings of Arkhangelsky et al. (2021), we include both types of weights, whereas in a robustness check we restrict the time weights to be equal. Should such a restriction be warranted, our estimate for the treatment effect should be more precise in this restricted version.
[9] Specifically, Arkhangelsky et al. (2019) show in simulations how SynthDID outperforms SC (both with and without penalized weights) in terms of both RMSE and bias in most cases (Tables 1, 2, 4 and 5). Next to SynthDID, other approaches have been introduced that relax some of the restrictions of DID and SC. A comparison of all these approaches is beyond the scope of this paper and we refer to Li and Shankar (2023) and Li and Van den Bulte (2023) for an excellent discussion on some of these approaches.

This assumption "rules out the possibility that a group gets treated because it experiences negative shocks" (De Chaisemartin and d'Haultfoeuille 2020, p. 2968). As discussed in the "Setting and Institutional Details" section, there is no coordination across municipalities that would make this assumption untenable. We next discuss the SynthDID approach in more detail, which we use in our focal analysis. In the "Results" section, we show the results of the DID method as a robustness check.[10]

TABLE 3: COMPARING DID, SYNTHETIC CONTROL, AND SYNTHDID

	Control Units	**Pre-Treatment Time Periods**
DID	Equally weighted but allowing for a baseline difference between treated and control units	Equally weighted but allowing for a baseline difference between pre-treatment and post-treatment periods
Synthetic Control	Weights chosen such that the weighted average pre-treatment outcome of the controls is approximately equal to the average pre-treatment outcome of the treated units	Zero weighted
SynthDID	Similarly weighted as in SC but regularized to increase the dispersion of the weights and allowing for a baseline difference between treated and control units	Weights chosen such that the weighted average of the pre-treatment outcome of the controls is approximately equal to the post-treatment outcome of the same controls, allowing for a baseline difference between pre-treatment and post-treatment periods

Synthetic Difference-in-Differences

The identification strategy relies on a stacked, SynthDID approach (for similar approaches see Berman and Israeli 2022 and Li et al. 2022), which not only can handle staggered introductions but also minimizes pre-treatment differences in the relevant grocery shopping behavior between households in treated and control municipalities. As noted above, the SynthDID estimator constructs a synthetic control group that matches the pre-treatment outcomes of the treated households up to a constant. Importantly, matching on pre-treatment outcomes allows us to indirectly match on unobservable factors. That is, to the degree that these factors cause co-variation in pre-treatment shopping behavior between treated and control households, they are implicitly included in the weighting approach (for a similar reasoning, see

[10] As discussed, the SC method does not regularize the weights on the potential controls, resulting in potentially overfitting issues. We return to the SC method in the "Results" section.

Abadie 2021). In addition, as explained above, SynthDID also allows for time-varying weights, because some pre-treatment periods may be more relevant in constructing the synthetic control. That is, the time weights allow for a focus on a subset of pre-treatment time periods, such that the weighted average of pre-treatment periods predicts average post-treatment outcomes (up to a constant). In sum, what SynthDID does is to reweight control units and time periods and subsequently apply a DID. Next, we formally discuss the implementation of the staggered SynthDID for our setting. The staggered introduction of the ban requires careful computation and aggregation of treatment effects (such as to a common treatment time), to avoid making comparisons between units that are both already treated (Roth et al. 2023). Thus, to analyze the multiple introductions of the ban at different times, we follow Berman and Israeli (2022) and Li et al. (2022) and determine the effect of the ban per cohort (r) of introductions separately and then aggregate them into an overall treatment effect (Arkhangelsky et al. 2021, p. 4114). Formally, to evaluate the effect on grocery behavior Y_{itr} of household i in week t in cohort r we use:

$$(\hat{\tau}_r, \hat{\alpha}_{0r}, \hat{\alpha}_{ir}, \hat{\gamma}_{tr}) = \underset{\tau_r, \alpha_{0r}, \alpha_{ir}, \gamma_{tr}}{\arg\min} \left\{ \sum_{i \in N_r} \sum_{t=\ell_{r,min}}^{\ell_{r,max}} (Y_{itr} - \alpha_{0r} - \alpha_{ir} - \gamma_{tr} - \text{ban\_post}_{itr} \cdot \tau_r)^2 \hat{\omega}_{ir} \hat{\lambda}_{tr} \right\}, \quad (1)$$

where $\hat{\tau}_r$ is the estimated effect of the introduction of a ban in cohort r; α_{0r} is an intercept; α_{ir} and γ_{tr} are household and week fixed effects, respectively; N_r is the set of eligible households in both the treated and control municipalities for a given cohort for which we observe a balanced panel; $\ell_{r,min}$ denotes the start of the analysis window for cohort r, whereas $\ell_{r,max}$ is the end; ban_post_{itr} is an indicator of whether household i in cohort r resides in a municipality in which the ban was active during period t; $\hat{\omega}_{ir}$ and $\hat{\lambda}_{tr}$ are unit weights for the control households and time weights, respectively, where the unit weights are chosen such that the pre-treatment period of the controls has a parallel trend to the households in treated municipalities and the time weights are chosen such that the pre- and post-period for the controls are similar (up to a constant). As in Arkhangelsky et al. (2021), the remaining idiosyncratic error is assumed to be Gaussian. We compute the overall treatment effect ($\hat{\tau}$), as well as its standard error using the treated units in all cohorts following the approach outlined in Clarke et al. (2023). Before turning to the estimation results, we discuss the sample used for each cohort.

Estimation Sample (Treated and Control Households)

The total pool of eligible households for our estimation sample contains all households for which we have complete information (e.g., municipality in which they reside). The treatment group for each cohort consists of all households residing in a municipality that introduced a ban at that time. We begin with an intent-to-treat (ITT) estimate, where we include all households within a treated municipality, regardless of whether they opt in. An ITT is a useful measure to understand the effect of the introduction of a ban at the population level (cf. the treatment of not receiving flyers anymore unless they opt in) as this is the level of implementation and interest for policymakers (Imbens and Rubin 2015, p. 514).[11,12] Moreover, ITTs are advised in the presence of incomplete data on treatment compliance (Gupta 2011), which is the case for households' opt-in decision. To estimate the effect of not receiving unsolicited flyers anymore, we also consider a local average treatment effect (LATE) estimate for households that did not opt in and stopped receiving flyers.[13] We report these in the "Subgroup Analysis" section.

In terms of the control households, we consider all households as potential controls for each cohort, as long as the municipality in which they are located has not yet been treated. The degree to which any such household is relevant as a control for each treatment cohort is determined by the SynthDID procedure, which allocates weights according to the pre-treatment period similarity between treatment and control households on the dependent variables. Table 4 specifies the timeline of the introduction of bans across cohorts and households serving as controls for each.

---

[11] The FDA, for example, requests ITT estimates for approval of new drugs (Pearl 2000, p. 261).
[12] We mainly expect an effect for households that do not opt in and stop receiving store flyers. Including households for which the situation did not change may therefore dilute the effects. Nevertheless, other households may have adjusted their grocery shopping behavior too. For instance, previous research shows that opting in tends to make consumers more responsive to marketing communications (Godinho de Matos and Adjerid 2022). We explore this heterogeneity in a subgroup analysis where we report results separately for each set of households depending on their opt-in decision.
[13] In the terminology of Imbens and Rubin (2015, p. 514), our ITT estimate reflects the *assignment to treatment* (i.e., default change), while our LATE estimate reflects the *receipt of treatment* (i.e., not receiving flyers anymore).

TABLE 4: BANS ACROSS COHORTS AND TREATED AND CONTROL GROUPS

Coh.	Municipality	Introduction	2017	2018	2019	2020	2021	N Treated	N Control
1	Amsterdam	January 2018						174	7,336
2	Haarlem	January 2020							
	Tilburg	January 2020						293	7,001
	Utrecht	January 2020							
3	The Hague	July 2020						183	6,027
4	Leiden	October 2020						29	5,971
5	Rotterdam	November 2020						213	5,968
-	Other	-							

Notes: coh. = cohort.

= Pre-ban period = Post-ban period = Households in municipality eligible as control

For each of the cohorts, we use an analysis window of ±52 weeks around the introduction of the ban to ensure pre-treatment similarity encompasses any seasonal variations (see also Table 4).[14] We only retain households in each cohort that are part of the household panel for the entire period. Furthermore, this analysis requires a balanced panel (i.e., no gaps or left- or right-censoring). Although for some dependent variables (e.g., grocery expenditure), we do not require the household to have made a trip to a store in all weeks to remain in the panel, for other dependent variables (e.g., share-based measures), we drop households for which we do not observe a trip in all weeks considered. Table 5 reveals the total number of households included: The total number of observations we used to estimate our model equals 3,452,280 for dependent variables unconditional on a trip and 1,438,840 for dependent variables conditional on a trip. Web Appendix B contrasts average household characteristics of treated and control households, both overall and by cohort.

Web Appendix C shows the pre-treatment outcomes for the SynthDID control and the treated group for all cohorts, which indicates that seasonality is similar for the SynthDID control and the treated group prior to the implementation of each ban. We also provide figures showing the average pre-treatment evolution of the raw (i.e., unweighted) data series. To contrast the quality of the SynthDID control vs. the raw data vis-à-vis the treated group, Table 6 shows the average pre-treatment correlations

[14] In robustness checks, we set the window to one quarter (±13 weeks) and half a year (± 26 weeks), which capture more short- and medium-term effects. We find substantively the same results across all three windows.

by dependent variable across cohorts. On average, the outcomes of the SynthDID control are more highly correlated with the outcomes of the treated group (.86) than the raw data (.64) in the pre-treatment period, but not so high as to cause concerns of overfitting.

TABLE 5: NUMBER OF HOUSEHOLDS AND OBSERVATIONS USED IN SYNTHDID

	Number of Households			Number of Observations
	Treated	**Control**	**Total**	
***Unconditional* on trip observed in a week**				
Total grocery expenditure Total number of units purchased Number of shopping trips Grocery expenditure at primary chain Number of shopping trips to primary chain	892	32,303	33,195	3,452,280
***Conditional* on trip observed in a week**				
Share of private labels Share of purchases on promotion Number of different retail chains visited Trip share HDs vs. traditional retailers	361	13,474	13,835	1,438,840

Notes: The total number of eligible households consists of all 15,486 households for which we have background information (e.g., municipality in which they reside). Of these households, 1,883 reside in a treatment municipality and 13,603 in a control municipality. A household can used as a control for multiple cohorts as long as it is not (or not yet) exposed to any treatment itself. The number of households in this table reflects the treated and control households that are part of the panel for the entire 104-week period. The number of observations reflects the 104 weeks of data that we consider for each household. HD = hard discounter.

TABLE 6: PRE-TREATMENT CORRELATION TREATED AND CONTROL

	Raw	**SynthDID**
Total grocery expenditure	.87	.94
Total number of units purchased	.82	.91
Share of private label	.37	.78
Share of purchases on promotion	.64	.82
Number of shopping trips	.71	.87
Number of different retail chains visited	.51	.82
Trip share HDs vs. traditional retailers	.17	.77
Grocery expenditure at primary chain	.81	.92
Number of shopping trips to primary chain	.83	.92
Average	.64	.86

Notes: Correlations are averaged across cohorts. For SynthDID, the control households are weighted as described in the "Synthetic Difference-in-Differences" section.

RESULTS

Table 7 reveals that the effect of a ban on unsolicited store flyers on total grocery expenditures is .40 euro, with a standard error of .43. The 95% confidence interval ranges from -.45 to 1.24 and includes zero. Irrespective of statistical significance, an effect of .40 indicates a change of less than 1% (i.e., mean spending across treated households prior to the ban was 55.08; .40/55.08 = .73%), with just .01 standard deviations (i.e., .40/40.58 = .01). That is, the effect appears, managerially and economically, unimportant. For the other dependent variables, we similarly find estimated parameters that are statistically not significant and economically small.[15] Following suggestions by McShane et al. (2024), Web Appendix D shows interval estimate curves for each dependent variable, for all levels from 0% to 99.9%. Importantly, across all dependent variables, the confidence interval includes zero across a wide range of values.

TABLE 7: SYNTHDID ESTIMATES OF THE IMPACT OF INTRODUCING A BAN

	Estimate		**95% Confidence Interval**		
	$\hat{\tau}$	**Jackknife SE**	**Lower Bound**	**Upper Bound**	***p*-Value**
Total grocery expenditure (in euros)	.40	.43	-.45	1.24	.36
Total number of units purchased	.09	.36	-.62	.80	.81
Share of private labels (%)	-.26	.38	-1.02	.49	.49
Share of purchases on promotion (%)	-.41	.26	-.92	.10	.12
Number of shopping trips	.01	.02	-.02	.05	.46
Number of different retail chains visited	.00	.01	-.03	.02	.88
Trip share HDs vs. traditional retailers (%)	-.05	.53	-1.09	.98	.92
Grocery expenditure at primary chain (in euros)	.19	.40	-.60	.98	.64
Number of shopping trips to primary chain	.00	.01	-.02	.03	.74

Note: HD = hard discounter. Standard errors are derived using the Jackknife method using the procedure outlined in Clarke et al. (2023). The bootstrap procedure leads to virtually similar standard errors.

To understand if the null results are driven by particular choices in our research design or a lack of power, we ran several robustness checks and conducted a power analysis. We first discuss and report the robustness checks. First, for outcome measures that are not expressed in shares, we estimate the same model for a log-transformed version of the dependent variables (i.e., total grocery expenditure, total number of units purchased, number of shopping trips, grocery expenditure at primary chain, and number of shopping trips to primary chain).[16] Second, estimating the model using a ±52-week window might

[15] Next to grocery expenditure and number of shopping trips to the primary chain, we also assessed changes in these variables at the secondary and tertiary retail chain and similarly did not find any significant changes.

[16] The choice of functional form is important (Goldfarb, Tucker, and Wang 2022), as applies for DID frameworks (Roth et al. 2023). Our SynthDID approach maximizes the degree to which the treatment and control households have parallel trends in the pre-treatment period. The assumptions that such parallel trends hold for the monotonic transformation of the dependent variable are relatively strict (Roth and Sant'Anna 2023). For the transformed dependent variables, we estimate another SynthDID regression that creates a control group based on the transformed, rather than original, dependent variables, again increasing similarity in pre-treatment trends.

restrict the sample and decrease statistical power. [17] Therefore, we re-estimate our SynthDID model using ±26-week and ±13-week windows and including all households for which we observe panel membership in these shorter windows, such that we expand the set of eligible households.[18] Third, considering that the ban was implemented in a staggered way, regulators may have paid special attention to the first ban, in Amsterdam in 2018, as it predates the other implementations by at least two years. Thereafter, regulators in other municipalities may have adjusted the way they implemented the subsequent bans. We recompute the estimates for Amsterdam only, as well as for all other municipalities, dropping Amsterdam. Fourth, to understand if our SynthDID estimator yields different results from a staggered two-way fixed effects DID estimator, we estimate the main model using this method. We also re-estimate the SynthDID model restricting the time weights to be equal, which should result in more precise estimates in case such a restriction is warranted.[19]

Table 8 contains the parameter estimates for all robustness checks. Log-transforming the measures does not alter the findings, as all estimates are not significant and economically small. Changing the analysis window to ±26 or ±13 weeks does not substantially alter our findings either. The parameters are comparable in size, and continue to fail to reach significance, despite the increase in the number of households in the sample. For this overlapping but slightly different sample, only the number of different retail chains visited appears to decrease somewhat in the analysis with a ±13-week window, yet the effect is economically negligible (.03 fewer retail chains per week, i.e., it would take more than 33 weeks for a

[17] In our main analyses for each cohort, we only include people who do not leave the panel throughout the estimation window (±52 weeks). Thus, our analysis is based on households who stay in the panel for at least two years. We also run our analyses on households who stay in the panel for the entire period (2017-2021) and find results that are substantially the same: For eight out of nine measures the treatment effects are not significant. Only for the share of purchases of promotion, do we find a small decrease of .74%-points ($\hat{\tau} = -.74$, $p = .02$). However, we note that by conditioning on households that stay in the panel for the entire period, older households (which are more likely to stay in the panel) make up a larger share of the sample. Our treatment heterogeneity analysis reported in a next section shows an effect on the share of purchases on promotion for this group, and their effect thus gets a larger weight in the overall treatment effect. Regardless, economically this change of less than one percentage point is small compared with the average share of purchases on promotion in the pre-treatment period of 21.6%.
[18] Using a time window of ±26 weeks (±13 weeks) expands the set of treated and control households to 1,000 (1,060) and 37,088 (41,149), respectively, for variables unconditional on a trip observed in a week, and to 558 (713) and 21,496 (29,460), respectively, for variables conditional on a trip observed in a week.
[19] As we discuss in Web Appendix E, the SC method leads to overfitting in our setting and is not considered as an alternative method here.

TABLE 8: ROBUSTNESS OF FINDINGS

	Focal	Ln(Y)	Time Windows: ±26 weeks	Time Windows: ±13 weeks	Specific Cohorts: Coh. 1 only	Specific Cohorts: Without coh. 1	DID	Equal time weights
Total grocery expenditure (in euros)	.40 [-.45, 1.24]	.03 [-.07, .13]	.34 [-.50, 1.17]	.21 [-.70, 1.12]	-.43 [-1.77, .91]	.60 [-.40, 1.60]	.70 [-.20, 1.59]	.56 [-.32, 1.45]
Total number of units purchased	.09 [-.62, .80]	.01 [-.02, .04]	.01 [-.69, .71]	.09 [-.64, .82]	-.38 [-1.57, .82]	.20 [-.63, 1.03]	.39 [-.37, 1.15]	.28 [-.47, 1.02]
Share of private labels (%)	-.26 [-1.02, .49]	-	-.36 [-.96, .25]	-.19 [-.80, .42]	-1.14 [-2.59, .31]	-.01 [-.89, .86]	-.28 [-1.05, .48]	-.29 [-1.06, .48]
Share of purchases on promotion (%)	-.41 [-.92, .10]	-	-.33 [-.79, .12]	-.13 [-.64, .39]	.00 [-1.03, 1.04]	-.53 [-1.11, .06]	-.42 [-.94, .09]	-.40 [-.91, .10]
Number of shopping trips	.01 [-.02, .05]	.01 [-.01, .02]	.00 [-.03, .04]	.00 [-.03, .04]	.01 [-.05, .07]	.01 [-.03, .06]	.02 [-.02, .06]	.02 [-.02, .06]
Number of different retail chains visited	.00 [-.03, .02]	.00 [-.01, .01]	-.02 [-.04, .00]	**-.03 [-.05, -.01]**	.04 [-.01, .09]	-.01 [-.05, .02]	.00 [-.03, .03]	.00 [-.03, .03]
Trip share HDs vs. traditional retailers (%)	-.05 [-1.09, .98]	-	-.05 [-.80, .71]	.16 [-.53, .85]	-1.11 [-3.28, 1.07]	.25 [-.93, 1.43]	-.22 [-1.26, .82]	-.16 [-1.21, .90]
Grocery expenditure at primary chain (in euros)	.19 [-.60, .98]	-.04 [-.18, .11]	.44 [-.37, 1.25]	-.05 [-.89, .78]	.25 [-.91, 1.42]	.17 [-.77, 1.11]	.55 [-.30, 1.41]	.37 [-.48, 1.22]
Number of shopping trips to primary chain	.00 [-.02, .03]	.00 [-.01, .01]	.00 [-.02, .03]	-.01 [-.03, .02]	.00 [-.05, .05]	.01 [-.02, .04]	.01 [-.02, .04]	.01 [-.02, .04]

Notes: 95% confidence interval in brackets; confidence intervals excluding zero in bold. Standard errors/confidence intervals for SynthDID are derived using the Jackknife procedure. For DID, we use two-way clustering by household and week (Berman and Israeli 2022). HD = hard discounter, coh. = cohort.

household to visit one retailer less). Notably, this small short-term effect goes to zero in the longer windows. The analysis that only includes Amsterdam (cohort 1) produces equally sized estimates, alsowithout any parameters reaching significance at traditional levels. The results for the analysis without Amsterdam again are small, with no significant parameters. Using the staggered two-way fixed effects DID approach yields comparable parameter estimates. The results of a more restrictive SynthDID model in which we restrict the time weights to be equal but retain the unit weights with regularization shows similar results as our focal model. Overall, the results of our robustness checks provide confidence in the validity of those obtained by our focal model.

Alternatively, null results can be driven by a lack of power to detect meaningful effects. To confirm that our sample size offers sufficient power to detect economically meaningful effects, we calculate a minimum detectable effect (MDE). Rather than assess the required sample size ("How many observations are required to show an effect of X?"), we consider the effect as a function of sample size (i.e., "How small is the smallest effect we can identify based on the sample size we have?"). In Web Appendix F, we detail how to calculate Minimum Detectable Effects (MDEs) given our setting. We find the MDEs, like the treatment effects, to be economically small (e.g., we should be able recover a change in expenditure of as little as 1.09 euro on a weekly basis, where our SynthDID and DID estimates are .40 and .71, respectively), suggesting enough power to find a meaningful effect should it be present. We provide all MDEs in Web Appendix F.

Subgroup Analysis

Thus far, our ITT analysis focused on the population level, i.e., the average change in behavior of all households in a municipality in which a ban is introduced, regardless of their decision to opt in or not. However, the effects can differ across households that opt in (and thus continue to receive store flyers) or households that already opted out prior to the ban (and thus continue to not receive flyers) on the one hand and households that do not opt in (and stop receiving such mail) on the other hand. To assess the effect of no longer receiving flyers, or LATE, we zoom in households for which implementing the municipality-level ban corresponded to such a change and contrast these to households for which no change in store flyer reception occurs. Formally, there are three subsamples for which we can analyze the impact of a ban: households that (1) previously received flyers and do not opt in ("compliers"), (2) previously received flyers and opt in to continue ("defiers"), and (3) already opted out prior to the ban and continue not to receive flyers ("always-takers").[20,21] The first group experiences a material change after the implementation (i.e., they no longer receive flyers); store flyer reception does not change for other two groups. Web Appendix H provides household demographics for each group; both overall and by cohort. For the subgroup analysis, we separately compare each subgroup with the control group from our main analysis and summarize the results in Table 9 (see Web Appendix I for the full results).[22] For defiers and always-takers, we find no statistically significant treatment effect across all dimensions, as expected. Beyond lacking statistical significance, the estimated effects also center around zero and are economically small.[23] This is in line with the proposed mechanism of search cost reduction and the heterogeneity across households herein: Defiers and always-takers are more extreme in their evaluation of the benefits they derive from receiving store flyers. For defiers, store flyers are likely to lower their search costs (to the extent that the benefits exceed the additional costs of opting in), such that they opt in to continue receiving flyers. The always-takers already indicated not wanting to receive store flyers and opted out before, and they continue not receiving them. Compliers, on the other hand, are affected by the ban. They did not opt out before but do not seem to value the benefits to outweigh the hassle or environmental costs,

---

[20] Our data contain information about the decision to (not) opt in for about 60% of the households in treatment municipalities. To include all households in the treatment municipalities, we use a random forest algorithm to impute the sticker usage of other households (see Web Appendix G). We obtain a hit rate of 94.8%; the results are virtually identical when we re-estimate our models using the 60% of households for which we have all information.
[21] In our analysis, we consider a households' opt-in decision in the year the ban was implemented. Our data shows that 95.1% of the households that opt in after implementation are still opted in one year later and 95.8% of those who did not opt in, continue to not opt in one year later. Nonetheless, to the extent that households may change their decision, we also redid the subgroup analysis using a ±13-week window and find similar results.
[22] The subgroups are defined only for treated municipalities, so we have two alternative ways to analyze subgroup effects in a SynthDID framework. The first is to re-estimate the weights $\widehat{\omega}_{ir}$ and $\widehat{\lambda}_{tr}$ for each subgroup, effectively altering the control group across the main analysis and subgroups, which would hamper comparability across the analyses. The second option runs a SynthDID for both subgroups using the weights $\widehat{\omega}_{ir}$ and $\widehat{\lambda}_{tr}$ generated in the main analysis. This option guarantees that the control group is the same, at the risk of having a control group pre-trend that may not optimally match the subgroup's pre-trend. In line with Berman and Israeli (2022), we select the latter option. The results of the former option are virtually identical.
[23] Combining defiers and always-takers into a single "non-complier" subgroup (for which there is no change in the reception of the story flyer) also does not change the results (see Web Appendix I).

or find the costs to contradict the new social norm too high, as they do not opt in after the legislation change. As such, they no longer learn about stores' advertised promotions through the store flyer. Interestingly though, also for compliers we find little change in any of the shopping measures. The results show a small decrease in the share of purchases of promotion ($\hat{\tau}$ = -.97, p = .03), but economically this change of less than one percentage point is small compared to the average share of purchases on promotion in the pre-treatment period (21.6%).[24]

TABLE 9: RESULTS SUBGROUP ANALYSES

	Compliers (without a sticker)	**Defiers (opted in)**	**Always-takers (opted out before)**
Total grocery expenditure (in euros)	.18	.70	.19
	[-1.11, 1.46]	[-.70, 2.10]	[-1.52, 1.90]
Total number of units purchased	-.08	-.03	.67
	[-1.10, .95]	[-1.28, 1.21]	[-.67, 2.01]
Share of private labels (%)	-.55	-.03	-.25
	[-1.91, .81]	[-.94, .88]	[-2.45, 1.94]
Share of purchases on promotion (%)	**-.97**	.27	-1.07
	[-1.82, -.11]	[-.45, .99]	[-2.33, .18]
Number of shopping trips	.01	.00	.05
	[-.05, .07]	[-.06, .05]	[-.03, .13]
Number of different retail chains visited	-.02	.01	.01
	[-.07, .02]	[-.03, .05]	[-.06, .09]
Trip share HDs vs. traditional retailers (%)	-1.04	.21	1.53
	[-3.06, .97]	[-.61, 1.04]	[-2.16, 5.21]
Grocery expenditure at primary chain	-.02	.79	-.66
	[-1.28, 1.23]	[-.43, 2.01]	[-2.40, 1.07]
Number of shopping trips to primary chain	.00	.01	.00
	[-.04, .04]	[-.02, .05]	[-.06, .06]

Notes: 95% confidence interval in brackets; confidence intervals excluding zero in bold. HD = hard discounter. Number of treated (control) households in the complier subgroup analysis: 344 (32,303) and 136 (13,474) for variables unconditional respectively conditional on a trip observed in a week; defier subgroup analysis: 372 (32,303) and 168 (13,474); always-takers subgroup analysis: 176 (32,303) and 57 (13,474).

[24] As an alternative to using all households as controls in the subgroup analysis, we can impose certain rules that would qualify the eligibility of each control. While we consider simple rules that would select the appropriate set of controls infeasible, we ran a robustness check in which we consider control households that receive flyers as controls for complier and defier subgroups, and those control households that already opted out as controls for the always-taker subgroup. The results are substantively similar: We again find no statistically significant treatment effect for defiers and always-takers, and a small decrease in the share of purchases on promotion for compliers ($\hat{\tau}$ = -.96, p = .03). In a second robustness check, we use the (to-be-treated) households from later cohorts that later also decide to not opt in as controls for the complier subgroup of the first cohort. For share of purchases on promotion, we find a similar effect size but with a larger p-value ($\hat{\tau}$ = -1.68, p = .11). The effects for the other eight shopping dimensions are of similar size and again not statically significant. Summarizing, using a subset of controls that are similar to the treated households before treatment or behave similarly when the treatment does happen yields substantively the same insights.

Treatment Heterogeneity

While we observe null effects across households, it might well be possible that some households were indeed affected by the change in regulation. For policymakers, it is relevant to know whether vulnerable segments (e.g., low-income, or older households) are affected by the ban and adjust their shopping behavior. To shed light on this question, we split treated and control households into different subgroups depending on households' (1) income (low, middle, high), (2) age in years (< 65, ≥ 65), (3) propensity to visit multiple retailers in pre-treatment period (≤ median, > median), and (4) price sensitivity (low, moderate, high[25]). For each of these subgroups, we re-estimate Equation 1 for each dependent variable and present the results in Web Appendix J.[26] Across the 90 sets of estimates (9 dependent variables x 10 household groups), we find only two significant effects: Older households and households that are moderately price sensitive (neither disagree not agree with the statement) appear to reduce their share of purchases on promotion ($\hat{\tau}$ = -.90, *p* = .03 and $\hat{\tau}$ = -1.28, *p* = .02 respectively). Yet, the reduction is small compared to the pre-treatment average of 21.6%; importantly, it does not seem to affect the households' overall expenditures. Interestingly, we find no significant effects for households that are highly price sensitive but only for those that are moderately price sensitive. One explanation could be that a larger extent of the highly price-sensitive households have opted in to continue receiving store flyers. Indeed, our data shows that about 39% of the highly price-sensitive households have opted in to continue to receive store flyers versus 32% of the moderately price-sensitive households.[27] We explore the relationship between household characteristics and opting in more formally in the "Profiling Households that (Do Not) Opt In to Receive Flyers" section. We first consider three possible options that

[25] Based on a question from GfK's annual household survey, "When buying groceries, I don't really pay attention to the price". We categorize the answer options "completely agree" and "agree" as "low", "neither agree nor disagree" as "moderate", and "disagree" and "completely disagree" as "high". In total, we have data for this question for 96% of all households (similar percentages for treated and control households).
[26] The results are robust to using all control households and the same weights $\hat{\omega}_{ir}$ and $\hat{\lambda}_{tr}$ as in the main analysis and to re-estimating Equation 1 using all control households (cf. same group control households, e.g., low income).
[27] At the same time, compared to younger households, a larger share of older households opt in (34% vs. 47%, respectively), while we still find a moderate effect on the share of purchases on promotion. This suggests that older households that do not opt in and stop receiving flyers were more likely to use flyers to guide their shopping behavior than younger households that do not opt in.

could explain a lack of meaningful changes in shopping behavior.

Explanations for a Lack of Meaningful Changes in Shopping Behavior

The preceding analyses show that banning unsolicited store flyers and switching to an opt-in system has no meaningful effect on household grocery shopping behavior, despite fears by retailers and flyer distributors that households may stick to the default option of not receiving store flyers because they find the process of opting in to be too much of a hassle or because the policy change changes what is socially desirable. As we discussed, the information provided in store flyers can guide consumers' shopping behavior by creating awareness for promoted products and lowering search costs. While households that strongly value store flyers are likely to have opted in to continue to receive store flyers and some households already opted out before the ban, such that little changes for these groups, there is a large group of households which are meaningfully affected by the ban in terms of receiving store flyers. No longer receiving them means they no longer hear about stores' promotions through the flyer, which may increase their search costs. Yet, our results show that even for those households, there is no substantive change in their shopping behavior. What could explain this lack of meaningful changes in shopping behavior? Below, we discuss three possible options.

First, in response to the ban, retailers may have adjusted their marketing strategy in the treatment municipalities. In the Netherlands, the content for store flyers is determined at the national level by all retailers, and most other marketing tools (e.g., TV and radio advertising) similarly target the entire country. As the implementation of the ban is staggered, adjustments to a retailers' national media efforts are unlikely. Still, any such national changes would influence households in our treatment and control markets similarly. Retailers may have created regionally customized billboards in treatment municipalities. We interviewed a general manager of the largest Dutch grocery franchise chain and a marketing director of JCDecaux, the largest billboard provider in the Netherlands, to understand retailers' possible reactions. The general manager signaled "a huge reliance on the store flyer" and also noted that costs saved due to the local bans had been redirected to other national advertising channels. The marketing director of the billboard provider identified "no change in sales" for billboard advertising in the treatment markets, explaining that billboards do not provide a perfect substitute for store flyers, due to their general usage to highlight a few products, typically of only one brand. As such, we consider local adjustments in marketing strategy an unlikely explanation for the lack of a meaningful effect.

Second, after households stop receiving store flyers, they may have adopted its digital version. Store flyers are available on retailers' websites and through apps, and several other websites aggregate flyers from all retailers. While many households consider digital flyers inferior to print flyers (Progressive Grocer 2024; Stantec 2024),[28] it is conceivable that compliers (who do not opt in to continue receiving print flyers) turn to digital versions. If so, households in treatment areas should display increased digital readership. To study this further, we use a single-wave survey, conducted in January 2021, just after the latest introduction of the ban, and obtain cross-sectional information from 2,186 households in treatment and control areas regarding their digital flyer readership (see Web Appendix K for the full details). We do not find evidence of higher readership of digital flyers in the treatment areas. Among households in the treatment area with the most recent ban implementation (i.e., Rotterdam in November 2020, two months prior to the survey), we also do not uncover any indicator that, among digital flyer users, they started doing so at different times than households from control areas. Thus, while some households may use digital flyers, the absence of a difference between treatment and control municipalities suggest that increased adoption in response to the ban is unlikely to explain the lack of any meaningful effect.

A final explanation is that the widespread dependence on store flyers may not have been as high as retailers and flyer distributors seem to believe. Previous literature already showed considerable heterogeneity in the usage of the store flyer, indicating that not all consumers who receive them, use them to decide on their shopping (Gázquez-Abad et al. 2014; van Lin and Gijsbrechts 2016), and that feature advertising only drives the store choice decision of certain consumers (Srinivasan and Bodapati 2006) and has only limited effects across the conversion funnel (Seiler and Yao 2017). Households that do not opt in

[28] Survey reports suggest that digital flyers are not perfect substitutes, as many consumers consider them inferior to the experience of reading a print flyer, potentially driven by the tactile experience of reading the print flyer (Progressive Grocer 2024; Stantec 2024). Such inferiority may lead to a lower engagement with its content.

and stop receiving store flyers after the ban, may not have used flyers in the first place. While they can no longer hear about stores' advertised promotions through the flyer and their search costs could have gone up, our results suggest that they did not derive many benefits from store flyers and did not rely on them to guide their shopping behavior. Put differently, the store flyer does not seem to have been a search cost shifter for these households. The concerns of retailers and flyer distributors that households stick to the default option of not receiving store flyers because they find the process of opting in to be too much of a hassle or because the policy change changes what is socially desirable, therefore seem ungrounded. While households that find store flyers worth their while likely opted in and continue to receive store flyers, households that derived little benefit from them and did not use them but did not opt out before either, no longer receive them. As such, it seems that the latter households actually stuck to the default of receiving store flyers because they found the process of opting out to be too much of a hassle or out of mere habit before the opt-in policy, such that they were fine with the change. In other words, the default change seems to have improved the sorting of households.

Summarizing, a lack of a meaningful effect could be due to adjustments in marketing strategy in treated municipalities, an increased usage of the digital store flyer, or because complying households do not use store flyers to guide their stopping behavior in any meaningful way. As we explain above, the first two explanations are unlikely such that we conclude that the 50% of households that no longer receive store flyers, were largely unresponsive to these costly marketing efforts, while those that actually find them worth their while opt in to keep receiving flyers. We seek a better profile of the potentially responsive households that opt in next.

Profiling Households that (Do Not) Opt In to Receive Flyers

Households that are prone to receive and read store flyers differ from other households on various dimensions, such as demographics and shopping-related characteristics (e.g., Gázquez-Abad et al. 2014; van Lin and Gijsbrechts 2016). Furthermore, research into people's general opt-in behaviors online (i.e., permission marketing) indicates that people who opt in differ from others in their demographics and online behavior (Kumar, Zhang, and Luo 2014). Therefore, to profile households that opt in to continue receiving store flyers after the ban (the “defier” subgroup), we use a binary logit model to regress their opt-in choice (1 = opt in, 0 = otherwise) on (1) household grocery shopping behavior 52 weeks before the ban (averages of total grocery expenditure, number of units purchased, share of private labels, share of purchases on promotion, number of shopping trips, number of different retail chains visited, trip share hard discounters vs. traditional retailers), (2) household demographics (household size, age of head of household, income class, social class), (3) household price sensitivity (based on their response to the statement “When buying groceries, I don’t really pay attention to the price”), (4) their readership of digital flyers (based on their response to the survey we previously reported), and (5) neighborhood characteristics obtained from Statistics Netherlands (percentage of households on social welfare, percentage of Western immigrants, percentage of non-Western immigrants, home ownership rate, average home value, population density, distance to the nearest grocery store, number of grocery stores in a 3-km radius, distance to the nearest department store, number of department stores in a 5-km radius). We do not include grocery expenditure and number of shopping trips to the primary chain because they cause multicollinearity due to their correlation with total grocery expenditure and number of shopping trips.

Table 10, column 2, contains the estimates. Not surprisingly, households that opt in (defiers) purchase a larger share of groceries on promotion ($\beta = 2.60$, $p < .01$), in line with findings on flyer readership by van Lin and Gijsbrechts (2016). This corroborates the face validity of our estimates. In addition, the remainder show a significant estimate for the average number of units purchased ($\beta = .02$, $p = .05$), such that households that purchase more units on a weekly basis also are more likely to opt in. In line with work on flyer readership by Gázquez-Abad et al. (2014) and van Lin and Gijsbrechts (2016), we also find that households that opt in are more likely to be older ($\beta = .24$, $p < .01$) and of larger size ($\beta = .25$, $p = .03$). They also come from lower social classes ($\beta = -.18$, $p = .03$) and reside in neighborhoods with a higher percentage of households on social welfare ($\beta = 5.06$, $p < .01$) and fewer grocery stores nearby ($\beta = -.03$, $p = .04$). Interestingly, readership of digital flyers does not correlate with opting in ($p = .59$), such that readers of digital flyers seem to be as inclined to continue to receive print flyers as those that do not read digital flyers. This is in line with survey reports suggesting that households consider

TABLE 10: LOGIT MODEL ESTIMATES OF OPT-IN CHOICE (N = 620)

	Defiers (opted in)	Compliers (without a sticker)	Always-takers (opted out before)
	Estimate (SE)	**Estimate (SE)**	**Estimate (SE)**
Household grocery shopping behavior in year before introduction of the ban			
Total grocery expenditure (in euros)	-.01 (.01)	-.01 (.01)	.03 (.01)**
Total number of units purchased	.02 (.01)*	.01 (.01)	-.06 (.01)**
Share of private labels (%)	.19 (.75)	-.71 (.68)	1.21 (.85)
Share of purchases on promotion (%)	2.60 (.93)**	-.70 (.87)	-2.82 (1.13)*
Number of shopping trips	-.04 (.11)	-.12 (.11)	.23 (.14)
Number of different retail chains visited	.32 (.23)	.04 (.22)	-.55 (.28)
Trip share HDs vs. traditional retailers (%)	.73 (.56)	.13 (.54)	-1.19 (.68)
Household demographics			
Household size[a]	.25 (.11)*	.11 (.10)	-.39 (.14)**
Age of head of household[b]	.24 (.06)**	.01 (.05)	-.22 (.06)**
Income class[c]	.01 (.02)	-.01 (.02)	.01 (.03)
Social class[d]	-.18 (.08)*	.08 (.08)	.11 (.10)
Household price sensitivity[e]	-.07 (.11)	-.08 (.10)	.22 (.13)
Reader of digital flyers	.10 (.19)	.00 (.18)	-.04 (.22)
Neighborhood characteristics			
Percentage of households on social welfare (%)	5.06 (1.58)**	-4.16 (1.57)**	-.49 (1.62)
Percentage of Western immigrants (%)	-.63 (1.00)	1.30 (.91)	-.77 (1.12)
Percentage of non-Western immigrants (%)	-.64 (.54)	.38 (.49)	.56 (.58)
Home-ownership rate (%)	.69 (.37)	-.19 (.34)	-.45 (.40)
Average home value (in 100,000 euros)	-.11 (.09)	-.05 (.08)	.21 (.10)*
Population density (1,000s of addresses per km^2)	.09 (.11)	.12 (.10)	-.24 (.12)*
Distance to the nearest grocery store (in kms)	.32 (.25)	.03 (.25)	-.56 (.29)
Number of grocery stores in a 3-km radius	-.03 (.02)*	.00 (.01)	.03 (.02)*
Distance to the nearest department store (in kms)	.01 (.11)	-.15 (.11)	.23 (.12)
Number of department stores in a 5-km radius	.08 (.05)	-.02 (.04)	-.07 (.05)
Intercept	-4.54 (1.06)**	.16 (.90)	1.47 (1.06)
Maximum variance inflation factor (VIF)	8.06	8.05	7.93

[a] 1 = 1 member, 2 = 2 members, 3 = 3 members, 4 = 4 members, 5 = 5 or more members.
[b] 1 = 12-19 y.o. (years old), 2 = 20-24 y.o., 3 = 25-29 y.o., 4 = 30-34 y.o., 5 = 35-39 y.o., 6 = 40-44 y.o., 7 = 45-49 y.o., 8 = 50-54 y.o., 9 = 55-64 y.o., 10 = 65-74 y.o., 11 = 75 y.o. or older.
[c] Net monthly income: 1 = <700, 2 = 700-900, 3 = 900-1,100, 4 = 1,100-1,300, 5 = 1,300-1,500, 6 = 1,500-1,700, 7 = 1,700-1,900, 8 = 1,900-2,100, 9 = 2,100-2,300, 10 = 2,300-2,500, 11 = 2,500-2,700, 12 = 2,700-2,900, 13 = 2,900-3,100, 14 = 3,100-3,300, 15 = 3,300-3,500, 16 = 3,500-3,700, 17 = 3,700-3,900, 18 = 3,900-4,100, 19 = >4,100.
[d] 1 = D (lower), 2 = C, 3 = B-minus, 4 = B-plus, 5 = A (upper); based on the education level and occupation of the head of the household.
[e] Based on a household's response to the statement "When buying groceries, I don't really pay attention to the price", with 1 = completely agree, 2 = agree, 3 = neither agree nor disagree, 4 = disagree, 5 = completely disagree.
* $p < .05$, ** $p < .01$.
Notes: The dependent variables are defined as 1 for households that opted in/without a sticker/already opted out before (respectively), 0 otherwise. The model is estimated on households' opt-in choice in the first year after the introduction of the ban, excluding households for which we impute their opt-in choice, households for which we do not observe a full year of shopping behavior before the introduction of the ban, and households for which we did not obtain their response to the statement on price sensitivity or readership of digital flyers. HD = hard discounter. Variables in percentages are divided by 100 for readability.

digital flyers inferior to print flyers (Progressive Grocer 2024; Stantec 2024), suggesting digital replacements may be more of a complement than a substitute. The above analysis presents retailers with a profile of potentially responsive households. However, to also understand which households do not put a sticker on their mailbox (the compliers) or already opted out (the always-takers), we present results from a similar analysis for these two groups in Table 10 (columns 3 and 4). Compliers seem less likely to reside in neighborhoods with a higher percentage of households on social welfare, but otherwise do not exhibit a discernible profile, possibly because they may be less extreme in their evaluation of the benefits of the store flyers. More interestingly, always-takers exhibit (for the most part) opposite characteristics to households that continue to receive store flyers in that they tend to spend more on a weekly basis but fewer units, purchase less on promotion, are of smaller size and younger age, reside in more affluent neighborhoods, and have access to more grocery stores nearby. All in all, these household profiles provide actionable insights for retailers in that it can help them understand the characteristics of those that prefer to continue to receive store flyer and those that do not want to receive flyers.

DISCUSSION

Despite the growth of digital media, retailers seem to believe that store flyers remain a critical medium to inform consumers about advantageous price discounts and to keep their offerings at the top of consumers' minds. But if the expensive mailers, which require paper, ink, and logistics resources as well, get thrown away unread, they serve little business purpose and are also environmentally problematic. We study a ban on unsolicited store flyers, by moving from an opt-out to an opt-in policy. Specifically, we investigate the extent to which such a change affects households' grocery shopping behavior. The policy change resulted in a reduction of store flyer distribution by about 50%. Contrary to concerns by retailers, the results of a stacked, synthetic difference-in-differences approach indicate that households' grocery shopping behavior did not meaningfully change. The findings are robust to different time windows and modeling approaches. Additional purpose-designed survey data corroborate the notion that shifting to digital store flyer usage is an unlikely driver of these null effects. We also do not find evidence that retailers adjusted their marketing communication in treated municipalities in response to a ban. Zooming in on the compliers, who were directly affected by the change in store flyer distribution and may face higher search costs, we find no evidence of a substantial change in grocery shopping behavior either.

Are there truly no changes, then? Across all analyses and robustness checks, a very small fraction of estimates are statistically different from zero, pointing to a shift in shopping behavior.[29] For instance, our robustness check using a 13-week time window shows a small drop in the number of retail chains visited (i.e., .03 fewer retail chains per week), yet the effect disappears when using longer time windows. This may be suggestive of minor short-term effects in store patronage, yet we note that the effect is economically speaking very small. Further, in some of the subgroup analyses, we observe a small reduction in the share of purchases on promotion. Similarly, the magnitude of change of 1%-point is economically limited. Finally, our results also do not indicate that vulnerable segments or any other household group are differentially affected by the legislative change. We observe a 1%-point reduction in the share of purchases of promotion among older and moderately price-sensitive households, but the effect is again negligible. Importantly, our results show that their overall expenditures remain unchanged.

Overall, our results suggest a win–win–win scenario for retailers, consumers, and the environment. We thus hope our study proves useful for policymakers writing legislation pertaining to physical advertising mail and retailers that rely on them, while also stimulating further research at this important intersection of marketing and public policy. Our findings offer several contributions and insights, particularly for retail managers who, even without a ban or policy change, might wonder whom to target with store flyers. According to our results, around 50% of flyer distribution can be cut without any adverse effects on household expenditures, in line with the old advertising adage: "Half the money I spend on advertising is wasted; the trouble is, I don't know which." The profiling analysis suggests that household characteristics that correlate with reading flyers (e.g., promotion sensitivity) also correlate with

[29] A total of nine out of 212 estimates (4.2%) reported in Tables 7-9 and in the Web Appendix attain statistical significance at $p < .05$. Out of the nine significant estimates, five are obtained using the robustness check using a synthetic control method, which shows signs of overfitting (see Web Appendix E). Ignoring the SC method, we obtain four significant estimates out of 203 (equaling 2.0%).

opting in to receiving flyers, rendering concerns that households stick to the default option of not receiving flyers unwarranted. Instead, our results seem to suggest that the share of households that lose measurable benefits due to the various costs (opportunity, environmental, and violating social norms) of opting-in is limited. As such, the ban may serve as a self-selection mechanism and improve the sorting of households.[30] The results of our analysis further suggest that a purpose-designed survey that measures flyer usage and the likelihood to opt in when considering the various costs involved can be used to gain ex-ante insights on the potential effect of a ban.

For retailers who have discontinued flyers unilaterally (and thus no longer distribute them to interested households either), such as OBI, the DIY market leader in Germany, and REWE the second-largest grocery retailer in Germany (Horizont 2022), our data also indicate that a sizeable segment of customers may still prefer to receive flyers. This is particularly worrisome as competitors can continue to distribute flyers without a ban. Anecdotal evidence supports such concerns. When Walmart temporarily stopped distributing its “Weekly Ad” in a handful of test markets in 2011 and 2012, store traffic fell (Kapner 2015a). It is worthwhile, however, to explicitly study such a one-sided withdrawal (by one retailer) and its effect on households’ shopping behavior vs. other retailers that continue to send flyers.

Our study also contributes to the academic literature on retailing and promotions. More specifically, we contribute to the literature on store flyers and feature advertising by being the first to show how banning unsolicited store flyers and moving from an opt-out to an opt-in policy affects grocery shopping behavior. As we show, even this drastic change does not affect shopping behavior, in line with Srinivasan and Bodapati’s (2006) and Seiler and Yao’s (2017) arguments that store flyers appeal only to a certain segment of consumers and that they have only limited effects across the conversion funnel.

Finally, our insights are relevant for policymakers considering legislation to curb the distribution

[30] It is likely that either pre- or post-ban, the sorting was worse. Should a change in flyer reception not lead to a change in shopping behavior, one could consider that as evidence that those households were incorrectly sorted prior to the ban (i.e., they received flyers but did not use them), and the change in default improves the sorting. Nevertheless, in the post-ban setting, sorting may still be imperfect, as households that have opted in may still be incorrectly sorted (e.g., they may opt in due to a fear of missing out, while in practice, their shopping behavior is unaffected by the store flyer).

of store flyers (or advertising mail more generally). Next to different municipalities of the Netherlands, an opt-in default has been introduced in parts of Australia, Canada, New Zealand. Other governments are considering a similar policy, but face strong opposition from the industry, including retailers. In response to such opposition, policymakers can point to our results, which indicate that such fears are unfounded with respect to households' shopping behavior and that an opt-in default can reduce waste without hurting their business. In terms of environmental savings, calculations based on inputs provided by the Association of Dutch Municipalities and flyer distributors show that the reduced distribution in the seven treatment municipalities alone (close to 1.5 million households in total) are similar to the annual energy consumption of approximately 1.7% for treated households (equal to 214,268 gigajoules) and .9% of the annual garbage (14,784 tons). While somewhat more modest, the reduced distribution also could lower emissions by .3% for the treated households (14,481 tons of CO_2) and decrease water usage by .4% (278,720 m^3) per year.[31] These environmental savings would roughly quadruple if the ban were rolled out nationally.[32] Although these savings might be smaller compared to some other sustainability initiatives, our results suggest that the savings from a ban on unsolicited store flyers can be obtained without meaningfully affecting consumers or retailers that depend on them. Yet, it is important to note that, next to environmental savings, governments considering similar legislation also need to take into account the downstream consequences on the distribution ecosystem, including job losses, when evaluating the societal benefits.[33] Importantly, the dominant revenue source for flyer distributors is tied to the number of prints (Stantec 2024). A cut in distribution by 50% will significantly reduce this number. The industry employs 20,000 delivery drivers alone and was sized close to 700 million euros in 2023 (Verbeek 2023) and the potential halving in the number of prints is likely to have a sizeable impact on jobs and the

[31] To facilitate policy evaluation for municipalities, the Association of Dutch Municipalities provided a tool ("Milieu-impacttool Ja/Ja-sticker") to assess environmental impacts, according to the opt-in rate and estimates of production costs and waste collection costs (VNG 2022).
[32] We predict households' opt-in choice in the other municipalities would they also introduce ban using a similar approach as we used to impute missing values in the treatment municipalities (see Web Appendix G), excluding the opt-in choices of geographically close households and municipality dummies.
[33] We thank the Editor for suggesting this broader account of the societal benefits.

industry in general. Indeed, the ban has not been rolled out nationwide as several municipalities share concerns about potential side effects (see also the "Setting and Institutional Details" section) (Stantec 2024).

We close by noting some limitations of our study, which suggest avenues for further research. Our study is the first to study the implementation of a ban on unsolicited store flyers. Reliance on these store flyers traditionally has been high in the Netherlands (Adformatie 2020), but store flyers are globally widespread. Our insights are promising for the many municipalities in the Netherlands that have not (yet) instituted such bans, but it is useful to determine if the observed null effects generalize across countries, some of which are in the process of implementing such changes (like France and Luxembourg). Second, we focus on grocery shopping contexts; other retailers, such as electronics and furniture retailers, which consumers visit less frequently and might not always be top-of-mind, also rely heavily on store flyers. Their flyers also feature promotions, but creating awareness of these less-frequently visited retailers represents another major purpose. We leave it to future research to test whether our results generalize to other retail industries. Third, policymakers (and flyer distributors and retailers alike) are keen to understand the effects of the ban, and to consider what implementation details might be revised. A national adoption of the ban may be on the horizon in the Netherlands as the Dutch government commissioned an exploratory report to study the legal and practical feasibility (Stantec 2024). Given the exploratory nature of this report, ample time remains until a potential adoption of a nationwide system. We encourage research into specific implementation details that might increase opt-in or out rates (and who continues to receive flyers) and thereby influence shopping behavior. Reception might differ substantially, depending on how the ban is communicated and implemented, including an outright ban, in which case interested households no longer receive flyers either. Such an outright ban would allow one to study the value of store flyers but requires either experimental geographical variation in applying these bans or sufficient exogenous variation in flyer reception at the individual level. Fourth, our results suggest that increased usage of digital flyers (e.g., apps) is no likely explanation for the lack of meaningful effects. Indeed, survey reports indicate that many consumers consider them inferior to print flyers

(Progressive Grocer 2024; Stantec 2024), which could lead to a lower engagement with its content. Still, digital flyers come in many forms and some retailers may be ahead in terms of digitization and provide a better experience. Future research could test how digital flyers should be presented to improve their experience and engagement.[34] All in all, we hope that our study proves useful for policymakers concerned with legislation surrounding physical advertising mail and retailers who rely on them and stimulates further research in the increasingly important intersection of marketing and public policy.

[34] To measure a retailer's level of digitization one could look at the richness of retailer apps in terms of the number of features. In a robustness check, we assessed whether shopping behavior at retailers with a lower versus higher digitization effort was differently affected but find no evidence for this. A retailer's digitization effort may be correlated with other retailer-specific factors however, and we encourage experimental research into this matter.

REFERENCES

Abadie, Alberto (2021), "Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects," *Journal of Economic Literature*, 59 (2), 391-425.

Adformatie (2020, November 11), "De folder is onmisbaar: zelfs de Tesla-rijder 42eft geen nee-sticker [The Store Circular Is Indispensable: Even a Tesla Driver Doesn't Have a "No" Sticker]," https://www.adformatie.nl/gedragsverandering/de-folder-onmisbaar.

Ailawadi, Kusum L. and Sunil Gupta (2014), "Sales Promotions," in *The History of Marketing Science,* Russell S. Winer and Scott A. Neslin, eds. Singapore/Hanover: World Scientific-Now Publishers, pp. 463-497.

Arkhangelsky, Dmitry, Susan Athey, David A. Hirshberg, Guido W. Imbens, and Stefan Wager (2019), "Synthetic Difference-in-Differences," working paper, https://www.aeaweb.org/conference/2020/preliminary/paper/Yt6FAHSa.

Arkhangelsky, Dmitry, Susan Athey, David A. Hirshberg, Guido W. Imbens, and Stefan Wager (2021), "Synthetic Difference-in-Differences," *American Economic* Review, 111 (12), 4088-4118.

Berman, Ron, and Ayelet Israeli (2022), "The Value of Descriptive Analytics: Evidence from Online Retailers," *Marketing Science*, 41 (6), 1074-1096.

Clarke, Damian, Daniel Pailañir, Susan Athey, and Guido Imbens (2023), "Synthetic Difference-in-Differences Estimation," IZA working paper no. 15907, https://docs.iza.org/dp15907.pdf.

De Chaisemartin, Clément, and Xavier d'Haultfoeuille (2020), "Two-Way Fixed Effects Estimators With Heterogeneous Treatment Effects." *American Economic Review*, 110 (9), 2964-2996.

Dinner, Isaac, Eric J. Johnson, Daniel G. Goldstein, and Kaiya Liu (2011), "Partitioning Default Effects: Why People Choose Not to Choose," *Journal of Experimental Psychology: Applied*, 17 (4), 332-341.

EUWID Pulp and Paper (2022, May 9), "Luxembourg's action on postal ads draws criticism from outside its borders," https://www.euwid-paper.com/news/markets/luxembourgs-action-on-postal-ads-draws-criticism-from-outside-its-borders-090522/.

Fernandez, Raquel, and Dani Rodrik (1991), "Resistance to Reform: Status Quo Bias in the Presence of Individual-Specific Uncertainty," *American Economic Review*, 81 (5), 1146-1155.

Fox, Edward J., and Stephen J. Hoch (2005), "Cherry-Picking," *Journal of Marketing*, 69 (1), 46-62.

Gabel, Sebastian, Dominik Molitor, and Martin Spann (2024), "Frontiers: The Effect of an Ad Ban on Retailer Sales: Insights from a Natural Experiment," *Marketing Science*, 43 (4), 723-733.

Gauri, Dinesh K., Brian Ratchford, Joseph Pancras, and Debabrata Talukdar (2017), "An Empirical Analysis of the Impact of Promotional Discounts on Store Performance," *Journal of Retailing*, 93 (3), 283-303.

----, K. Sudhir, and Debabrata Talukdar (2008), "The Temporal and Spatial Dimensions of Price Search: Insights from Matching Household Survey and Purchase Data," *Journal of Marketing Research*, 45 (2), 226-240.

Gázquez-Abad, Juan Carlos, Francisco J. Martínez-López, and Vanesa Barrales-Molina (2014), "Profiling the Flyer-Prone Consumer," *Journal of Retailing and Consumer Services*, 21 (6), 966-975.

----, and ---- (2016), "Understanding the Impact of Store Flyers on Purchase Behaviour: An Empirical Analysis in the Context of Spanish Households," *Journal of Retailing and Consumer Services*, 28 (1), 263-273.

Gedenk, Karen, Scott A. Neslin, and Kusum L. Ailawadi (2010), "Sales Promotion," in *Retailing in the 21st Century*, Manfred Krafft and Murali K. Mantrala, eds. Berlin/Heidelberg: Springer, pp. 393-407.

GfK/NMO (2024), "Bereiksonderzoek voor folders 2023/2024 [Reach Study for Flyers 2023/2024]," https://spotta.nl/resources/Pdf/Bereiksonderzoek%20Folders%202023%202024.pdf.

Gijsbrechts, Els, Katia Campo, and Tom Goossens (2003), "The Impact of Store Flyers on Store Traffic and Store Sales: A Geo-Marketing Approach," *Journal of Retailing*, 79 (1), 1-16.

Godinho de Matos, Miguel, and Idris Adjerid (2022), "Consumer Consent and Firm Targeting after GDPR: The Case of a Large Telecom Provider," *Management Science*, 68 (5), 3330-3378.

Goldfarb, Avi, Catherine Tucker, and Yanwen Wang (2022), "Conducting Research in Marketing with Quasi-Experiments," *Journal of Marketing*, 86 (3), 1-20.

Goodman-Bacon, Andrew (2021), "Difference-in-Differences with Variation in Treatment Timing," *Journal of Econometrics*, 225 (2), 254-277.

GroceryDrive (2018, January 10), "Study: Consumers prefer print to digital when it comes to ad circulars," https://www.grocerydive.com/news/grocery—study-consumers-prefer-print-to-digital-when-it-comes-to-ad-

circulars/534350/.
Gupta, Sandeep K. (2011), “Intention-to-Treat Concept: A Review,” *Perspectives in Clinical Research*, 2 (3), 109-112.
Guyt, Jonne, and Els Gijsbrechts (2018), “On Consumer Choice Patterns and the Net Impact of Feature Promotions,” *International Journal of Research in Marketing*, 35 (3), 490-508.
----, and ---- (2020), “Evaluating the Effectiveness of Retailer-Themed Super Saver Events,” *Journal of Marketing*, 84 (2), 92-113.
Haas, Michaela (2021, December 6), “Solutions: How Amsterdam Got Rid of Junk Mail,” *Atlanta Journal Constitution,* https://www.ajc.com/opinion/solutions-how-amsterdam-got-rid-of-junk-mail/AAR7VVYP7ZAA7MLTM4VYMPOOFE/.
Honka, Elisabeth, Ali Hortaçsu, and Matthijs Wildenbeest (2019), “Empirical Search and Consideration Sets,” in *Handbook of the Economics of Marketing*, Jean-Pierre Dubé and Peter E. Rossi, Eds., Amsterdam: North-Holland, pp. 193-257.
Horizont (2022, July 27), “Auch Rewe verzichtet künftig auf gedruckte Werbeprospekte [Rewe will also stop using printed advertising brochures], ” https://www.horizont.net/marketing/nachrichten/ab-2023-auch-rewe-verzichtet-kuenftig-auf-gedruckte-werbeprospekte-201577.
Imbens, Guido W. and Donald. B. Rubin (2015), “Causal Inference in Statistics, Social, and Biomedical Sciences,” New York City, NY: Cambridge University Press.
Johnson, Eric. J, and Daniel Goldstein (2003), “Do Defaults Save Lives?” *Science*, 302 (5649), 1338-1339.
Kapner, Suzanne (2015a, March 11), “Retailers Can’t Shake the Circular Habit,” *The Wall Street Journal*, https://www.wsj.com/articles/retailers-cant-shake-the-circular-habit-1426113760.
---- (2015b, August 18), “Retailers Cut Back on Newspaper Circulars,” *The Wall Street Journal*, https://www.wsj.com/articles/retailers-cut-back-on-newspaper-circulars-1439945338.
Kumar, Viswanathan, Xi Zhang, and Anita Luo (2014), “Modeling Customer Opt-in and Opt-out in a Permission-Based Marketing Context,” *Journal of Marketing Research*, 51 (4), 403-419.
Lemon, Katherine N., and Stephen M. Nowlis (2002), “Developing Synergies between Promotions and Brands in Different Price-Quality Tiers,” *Journal of Marketing Research*, 39 (2), 171-185.
Li, Kathleen T. (2020), “Statistical Inference for Average Treatment Effects Estimated by Synthetic Control Methods,” *Journal of the American Statistical Association*, 115 (532), 2068-2083.
----, and Venkatesh Shankar (2023), “A Two-Step Synthetic Control Approach for Estimating Causal Effects of Marketing Events,” *Management Science*, forthcoming.
----, and Garrett P. Sonnier (2023), “Statistical Inference for the Factor Model Approach to Estimate Causal Effects in Quasi-Experimental Settings,” *Journal of Marketing Research*, 60 (3), 449-472
----, and Christophe Van den Bulte (2023), “Augmented Difference-in-Differences,” *Marketing Science*, 42 (4), 746-767.
Li, Xingyi, Yiting Deng, Puneet Manchanda, and Bert De Reyk (2024), “Can Lower Expert Opinions Lead to Better Consumer Ratings?: The Case of Michelin Stars,” *Management Science*, forthcoming
Markteffect (2022), “JA-JA monitor,” report.
McShane, Blakeley B., Eric T. Bradlow, John G. Lynch, and Robert J. Meyer (2024), ““‘Statistical Significance” and Statistical Reporting: Moving Beyond Binary,” *Journal of Marketing*, 88 (3), 1-19.
Mettler, Suzanne, and Joe Soss (2004), “The Consequences of Public Policy for Democratic Citizenship: Bridging Policy Studies and Mass Politics,” *Perspectives on Politics*, 2 (1), 55-73.
Near (2022, June 22), ““‘Oui Pub” : Mailbox flyers will soon be “opt in” rather than “opt out” in France – how should retailers adapt their media strategy?,” https://near.com/blogs/oui-pub-mailbox-flyers-will-soon-be-opt-in-rather-than-opt-out-in-france-how-should-retailers-adapt-their-media-strategy/.
NOS (2021, October 8), “Ja-sticker beleeft opmars: 'Er is 27 miljoen kilo papier bespaard' [YES-Sticker on the Rise: 27 Million Kilo of Paper Has Been Saved],” https://nos.nl/artikel/2400839-ja-sticker-beleeft-opmars-er-is-27-miljoen-kilo-papier-bespaard.
Pearl, Judea (2000), “Causality – Models, Reasoning, and Inference,” Cambridge, UK: Cambridge University Press.
Pieters, Rik, Michel Wedel, and Jie Zhang (2007), “Optimal Feature Advertising Design Under Competitive

Clutter," *Marketing Science*, 53 (11), 1815-1828.
Progressive Grocer (2024, April 4), "Yes, Print Ads Are Still Relevant," https://progressivegrocer.com/yes-print-ads-are-still-relevant.
Retailtrends (2016, April 21), "Amsterdam stemt in met ja/ja-sticker [Amsterdam Approves Yes/Yes Sticker]," https://retailtrends.nl/news/44319/-amsterdam-stemt-in-met-ja-ja-sticker.
Roth, Jonathan, Pedro Sant'Anna, Alyssa Bilinski, and John Poe (2023), "What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature," *Journal of Econometrics*, 235 (2), 2218-2244.
----, and ---- (2023), "When Is Parallel Trends Sensitive to Functional Form?" *Econometrica*, 91 (2), 737-747.
Samuelson, William, and Richard Zeckhauser (1988), "Status Quo Bias in Decision Making," *Journal of Risk and Uncertainty*, 1, 7-59.
Schultz, P. Wesley (1999), "Changing Behavior With Normative Feedback Interventions: A Field Experiment on Curbside Recycling," *Basic and Applied Social Psychology*, 21 (1): 25-36.
Seiler, Stephan, and Song Yao (2017), "The Impact of Advertising Along the Conversion Funnel," *Quantitative Marketing and Economics*, 15 (3), 241-278.
Srinivasan, V. Seenu, and Anand V. Bodapati (2006), "The Impact of Feature Advertising on Customer Store Choice," working paper, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=901458.
Stantec (2024), "Verkenning Landelijke Invoering Opt-in Systeem voor Ongeadresseerd Reclamedruk [Exploration of Nationwide Implementation of an Opt-in System for Unsolicited Advertising Mail]," https://www.rijksoverheid.nl/documenten/rapporten/2024/06/05/bijlage-3-onderzoek-landelijke-invoering-opt-insysteem.
Tankard, Margaret E., and Elizabeth Levy Paluck (2016), "Norm Perception As a Vehicle for Social Change," *Social Issues and Policy Review*, 10 (1), 181-211.
Van Lin, Arjen and Els Gijsbrechts (2016), "The Battle for Health and Beauty: What Drives Supermarket and Drugstore Category-Promotion Lifts?" *International Journal of Research in Marketing*, 33 (3), 557-577.
Verbeek, Jan (2023, April 29), "Folderbezorger Spotta wil Nee-stickers omzeilen [Flyer distributor Spotta wants to bypass No-sticker]," *Financieel Dagblad*, https://fd.nl/bedrijfsleven/1474782/folderbezorger-spotta-wil-nee-stickers-omzeilen.
Vericast (2020, June 11), "Unprecedented Times Prove Retailer Circulars Remain a Powerful Engagement Tool," https://www.vericast.com/insights/blog/unprecedented-times-prove-retailer-circulars-remain-a-powerful-engagement-tool/.
VNG (2022), "Tool helpt bij besluit over stickers ongeadresseerd drukwerk [Tool Helps with Decision over Stickers for Unsolicited Advertising Mail]," https://vng.nl/nieuws/tool-helpt-bij-besluit-over-stickers-ongeadresseerd-drukwerk.
Widdecke, Kai A., Wiebke I.Y. Keller, Karen Gedenk, and Barbara Deleersnyder (2023), "Drivers of the Synergy between Price Cuts and Store Flyer Advertising at Supermarkets and Discounters," *International Journal of Research in Marketing*, 40 (2), 455-474.
Woolf, Amber-Leigh (2019, June 21), "Junk Mail and Windscreen Wiper Fliers Could Soon be Regulated in Wellington," https://www.stuff.co.nz/environment/113629701/junk-mail-and-windscreen-wiper-fliers-could-soon-be-regulated-in-wellington.

WEB APPENDICES

BANNING UNSOLICITED STORE FLYERS: DOES HELPING THE ENVIRONMENT HURT RETAILING?

Table of Contents

WEB APPENDIX A: EXAMPLES OF STICKERS USED TO OPT OUT/IN	**p. 2**
WEB APPENDIX B: COMPOSITION OF TREATED AND CONTROL SAMPLES	**p. 3**
WEB APPENDIX C: PRE-TREATMENT TRENDS	**p. 5**
WEB APPENDIX D: INTERVAL ESTIMATE CURVES	**p. 11**
WEB APPENDIX E: SYNTHETIC CONTROL METHOD	**p. 13**
WEB APPENDIX F: POWER TO DETECT MEANINGFUL EFFECTS	**p. 17**
WEB APPENDIX G: IMPUTATION OF OPT-IN CHOICE	**p. 20**
WEB APPENDIX H: COMPOSITION OF SUBGROUPS	**p. 21**
WEB APPENDIX I: RESULTS OF SUBGROUP ANALYSES	**p. 27**
WEB APPENDIX J: TREATMENT HETEROGENEITY	**p. 29**
WEB APPENDIX K: READERSHIP OF DIGITAL STORE FLYERS	**p. 35**

WEB APPENDIX A: EXAMPLES OF STICKERS USED TO OPT OUT/IN

FIGURE WA1: EXAMPLES OF STICKERS USED TO OPT OUT/IN

A: OPT-OUT APPROACH	B: OPT-IN APPROACH
	
Translation: No (“NEE”) to unsolicited advertising mail | Yes (“JA”) to community newspapers	Translation: Yes (“JA”) to unsolicited advertising mail | Yes (“JA”) to community newspapers

Note: In the opt-out approach, households receive store flyers and other unsolicited advertising mail unless they use the sticker on the left. In the opt-in approach, households receive no store flyers unless they use the sticker on the right.

WEB APPENDIX B: COMPOSITION OF TREATED AND CONTROL SAMPLES

Table WA1 show descriptive statistics of the household demographics for treated and control households across all cohorts. Table WA2 show similar statistics per cohort.

TABLE WA1: DEMOGRAPHICS TREATED AND CONTROL HOUSEHOLDS, ACROSS COHORTS

	Treated (N = 892)				Control (N = 32,303)			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Household size[a]	2.09	1.16	1	5	2.42	1.22	1	5
Age of head of household[b]	7.88	2.19	1	11	7.80	2.17	1	11
Income class[c]	9.70	5.07	1	19	10.25	4.92	1	19
Social class[d]	3.17	1.36	1	5	3.07	1.36	1	5

[a] 1 = 1 household member, 2 = 2 household members, 3 = 3 household members, 4 = 4 household members, 5 = 5 or more household members.
[b] 1 = 12-19 y.o. (years old), 2 = 20-24 y.o., 3 = 25-29 y.o., 4 = 30-34 y.o., 5 = 35-39 y.o., 6 = 40-44 y.o., 7 = 45-49 y.o., 8 = 50-54 y.o., 9 = 55-64 y.o., 10 = 65-74 y.o., 11 = 75 y.o. or older.
[c] Net income per month: 1 = below 700, 2 = 700-900, 3 = 900-1,100, 4 = 1,100-1,300, 5 = 1,300-1,500, 6 = 1,500-1,700, 7 = 1,700-1,900, 8 = 1,900-2,100, 9 = 2,100-2,300, 10 = 2,300-2,500, 11 = 2,500-2,700, 12 = 2,700-2,900, 13 = 2,900-3,100, 14 = 3,100-3,300, 15 = 3,300-3,500, 16 = 3,500-3,700, 17 = 3,700-3,900, 18 = 3,900-4,100, 19 = 4,100 or more.
[d] 1 = D (lower), 2 = C, 3 = B-minus, 4 = B-plus, 5 = A (upper); based on the education level and occupation of the head of the household.

TABLE WA2: DEMOGRAPHICS TREATED AND CONTROL HOUSEHOLDS, ACROSS COHORTS

A: COHORT 1

	Treated (N = 174)				Control (N = 7,336)			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Household size[a]	1.84	1.04	1	5	2.38	1.20	1	5
Age of head of household[b]	8.61	1.87	2	11	7.97	2.13	2	11
Income class[c]	8.89	5.12	1	19	9.84	4.74	1	19
Social class[d]	2.96	1.44	1	5	2.94	1.38	1	5

B: COHORT 2

	Treated (N = 293)				Control (N = 7,001)			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Household size[a]	2.23	1.21	1	5	2.42	1.22	1	5
Age of head of household[b]	7.62	2.31	1	11	7.79	2.20	1	11
Income class[c]	10.17	5.11	1	19	10.18	4.87	1	19
Social class[d]	3.28	1.32	1	5	3.05	1.36	1	5

C: COHORT 3

	Treated (N = 183)				Control (N = 6,027)			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Household size[a]	2.13	1.13	1	5	2.42	1.23	1	5
Age of head of household[b]	7.78	2.10	3	11	7.75	2.17	2	11
Income class[c]	9.95	5.30	1	19	10.43	5.01	1	19
Social class[d]	3.24	1.33		5	3.12	1.35	1	5

TABLE WA2 (CONTINUED)

D: COHORT 4

	Treated (N = 29)				Control (N = 5,971)			
	Mean	**SD**	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**
Household size[a]	2.14	1.25	1	5	2.43	1.23	1	5
Age of head of household[b]	7.41	2.28	3	11	7.74	2.17	2	11
Income class[c]	8.59	4.57	2	19	10.46	5.00	1	19
Social class[d]	3.21	1.30	1	5	3.13	1.35	1	5

E: COHORT 5

	Treated (N = 213)				Control (N = 5,968)			
	Mean	**SD**	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**
Household size[a]	2.08	1.16	1	5	2.43	1.23	1	5
Age of head of household[b]	7.81	2.19	2	11	7.73	2.17	2	11
Income class[c]	9.67	4.78	1	19	10.46	5.00	1	19
Social class[d]	3.10	1.38	1	5	3.13	1.35	1	5

[a] 1 = 1 household member, 2 = 2 household members, 3 = 3 household members, 4 = 4 household members, 5 = 5 or more household members.

[b] 1 = 12-19 y.o. (years old), 2 = 20-24 y.o., 3 = 25-29 y.o., 4 = 30-34 y.o., 5 = 35-39 y.o., 6 = 40-44 y.o., 7 = 45-49 y.o., 8 = 50-54 y.o., 9 = 55-64 y.o., 10 = 65-74 y.o., 11 = 75 y.o. or older.

[c] Net income per month: 1 = below 700, 2 = 700-900, 3 = 900-1,100, 4 = 1,100-1,300, 5 = 1,300-1,500, 6 = 1,500-1,700, 7 = 1,700-1,900, 8 = 1,900-2,100, 9 = 2,100-2,300, 10 = 2,300-2,500, 11 = 2,500-2,700, 12 = 2,700-2,900, 13 = 2,900-3,100, 14 = 3,100-3,300, 15 = 3,300-3,500, 16 = 3,500-3,700, 17 = 3,700-3,900, 18 = 3,900-4,100, 19 = 4,100 or more.

[d] 1 = D (lower), 2 = C, 3 = B-minus, 4 = B-plus, 5 = A (upper); based on the education level and occupation of the head of the household.

WEB APPENDIX C: PRE-TREATMENT TRENDS

FIGURE WA2: COHORT 1

A: TOTAL GROCERY EXPENDITURE | B: TOTAL NO. OF UNITS PURCHASED | C: % OF PRIVATE LABELS | D: % OF PURCHASES ON PROMOTION | E: NO. OF SHOPPING TRIPS

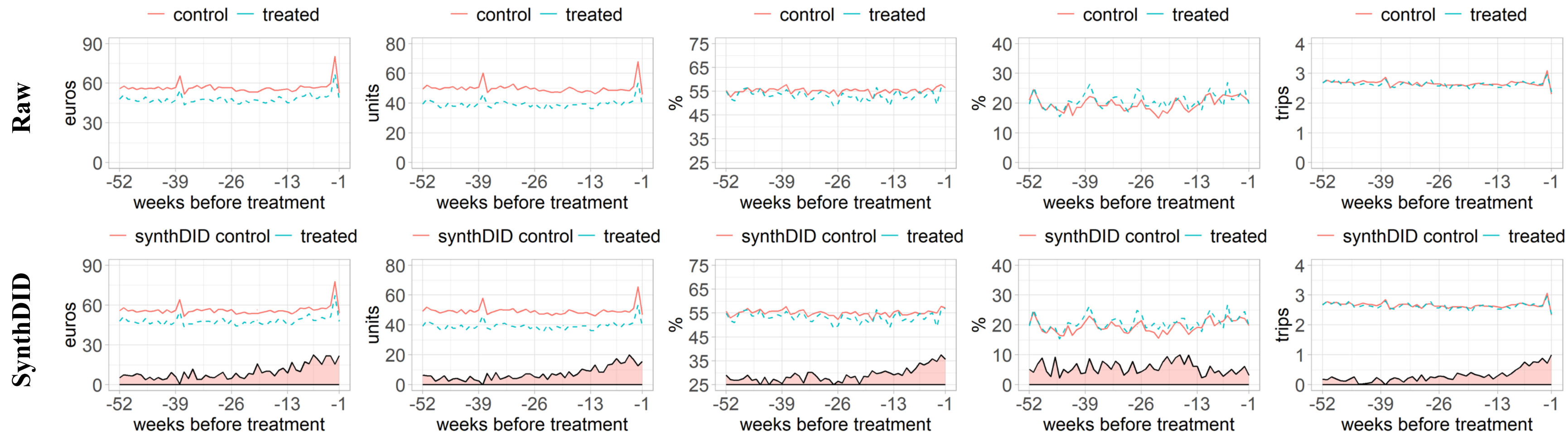
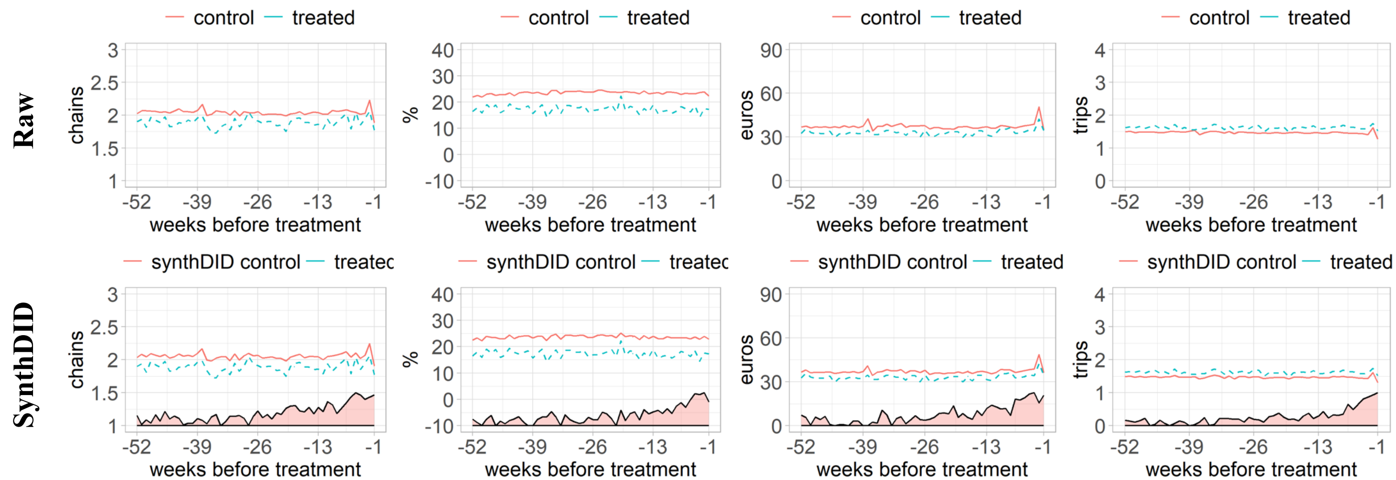


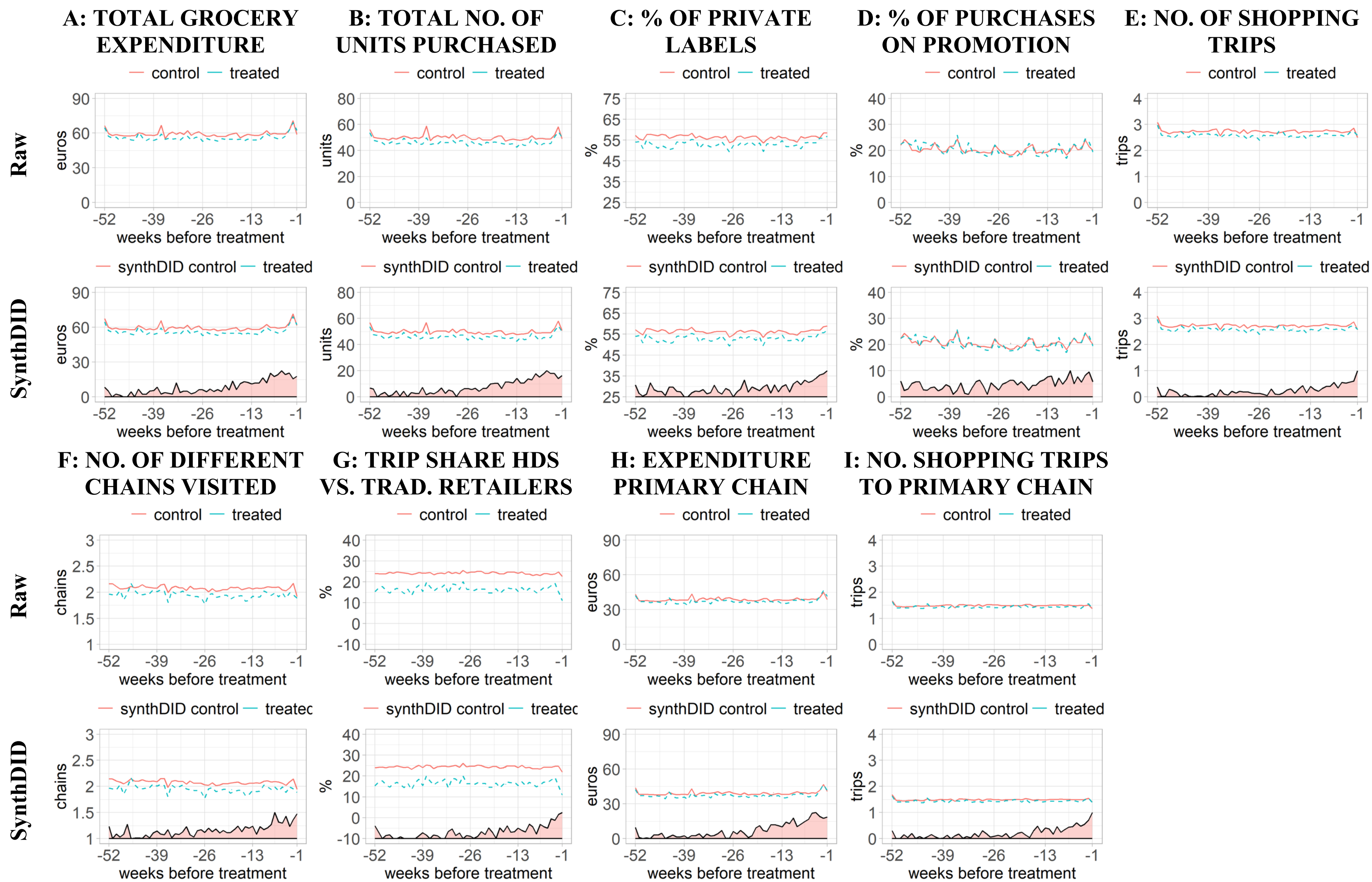
FIGURE WA2 (CONTINUED)

F: NO. OF DIFFERENT CHAINS VISITED | **G: TRIP SHARE HDS VS. TRAD. RETAILERS** | **H: EXPENDITURE PRIMARY CHAIN** | **I: NO. SHOPPING TRIPS TO PRIMARY CHAIN**



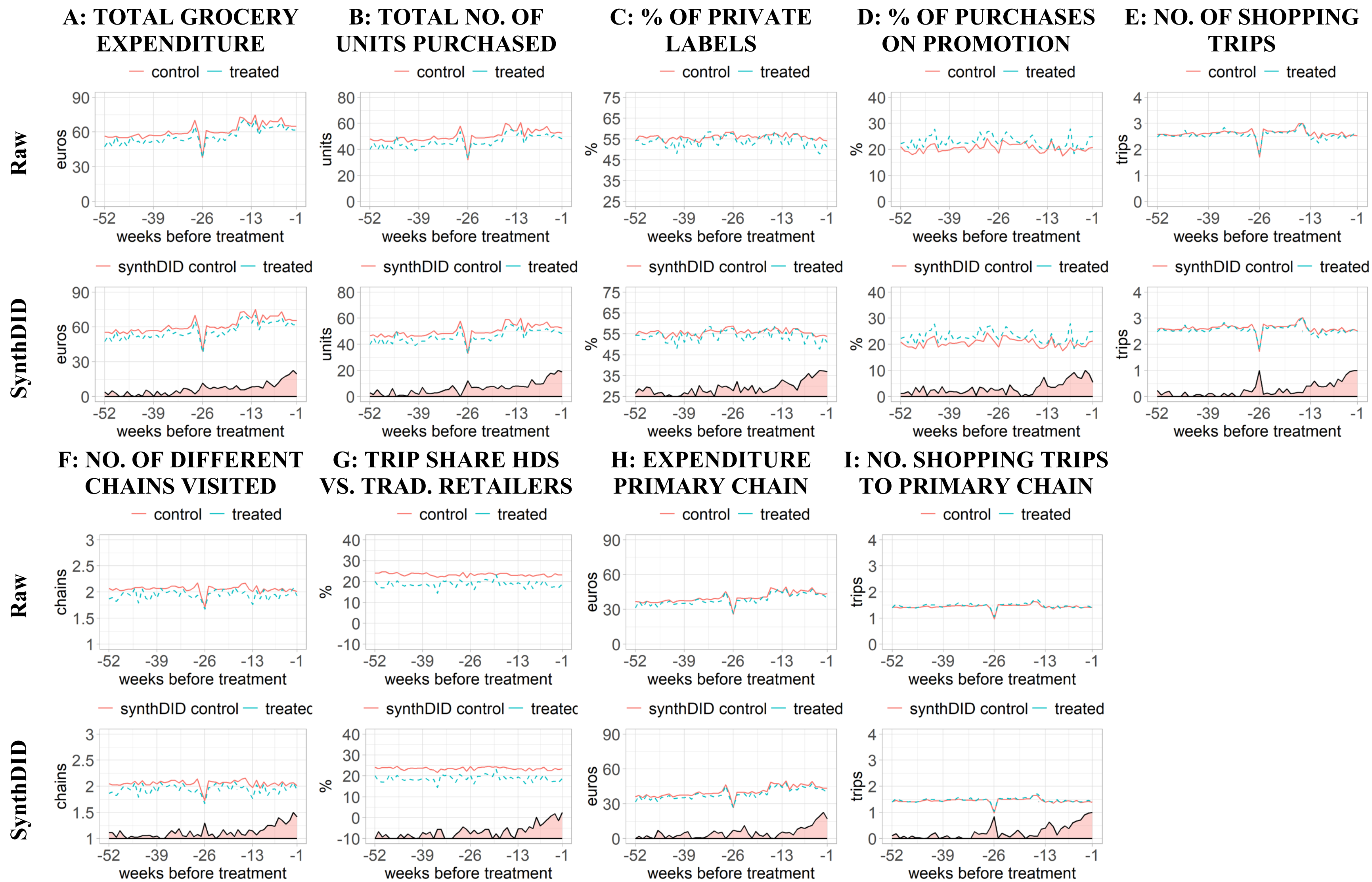
Notes: Dashed (solid) lines represent the weekly average for treated (control) households. The weighted average for the SynthDID control households is weighted as described in the “Synthetic Difference-in-Differences” section. The time weights used to average pre-treatment periods in the SynthDID approach are at the bottom of each plot.

FIGURE WA3: COHORT 2



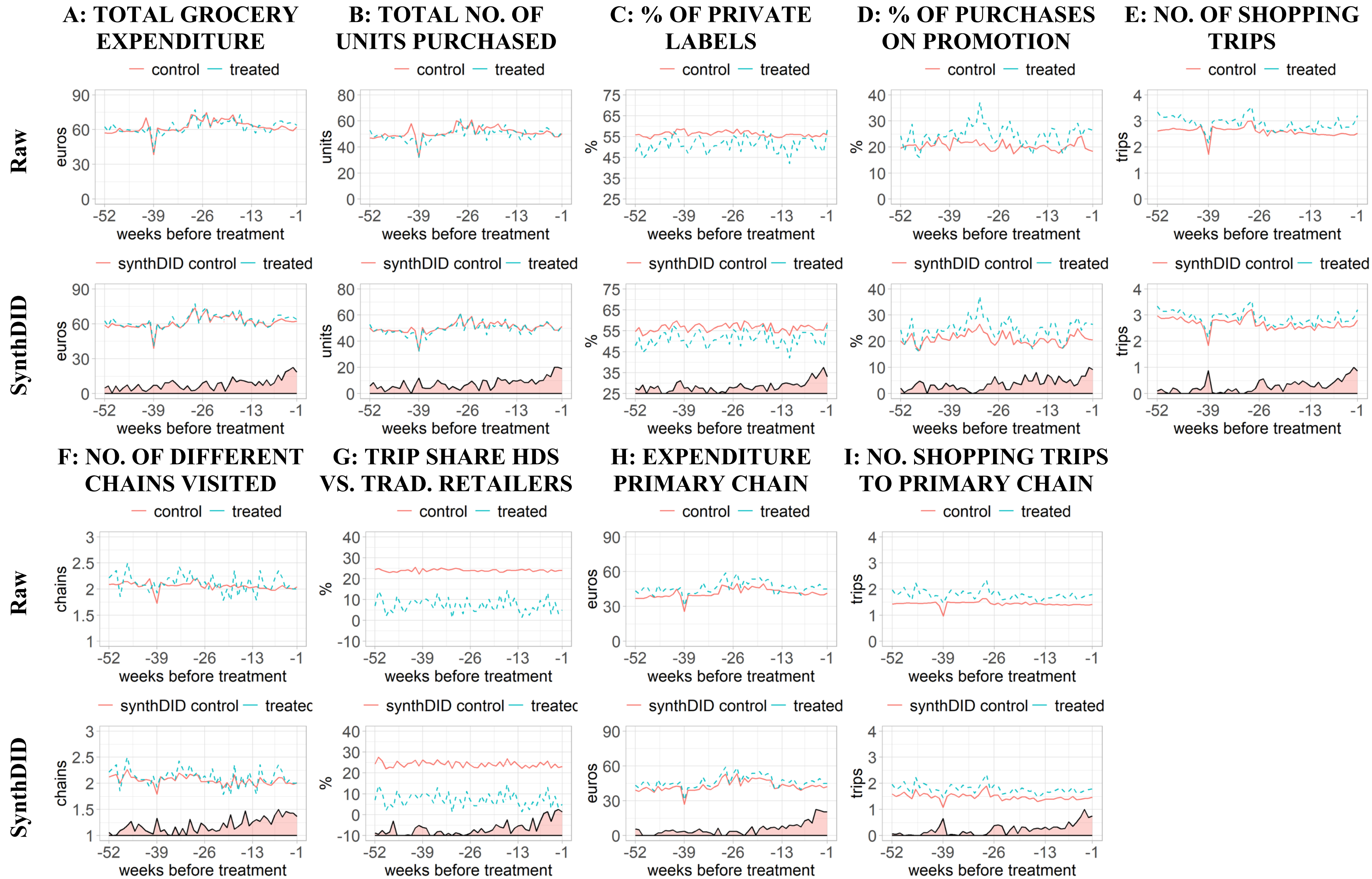
Notes: Dashed (solid) lines represent the weekly average for treated (control) households. The weighted average for the SynthDID control households is weighted as described in the "Synthetic Difference-in-Differences" section. The time weights used to average pre-treatment periods in the SynthDID approach are at the bottom of each plot.

FIGURE WA4: COHORT 3



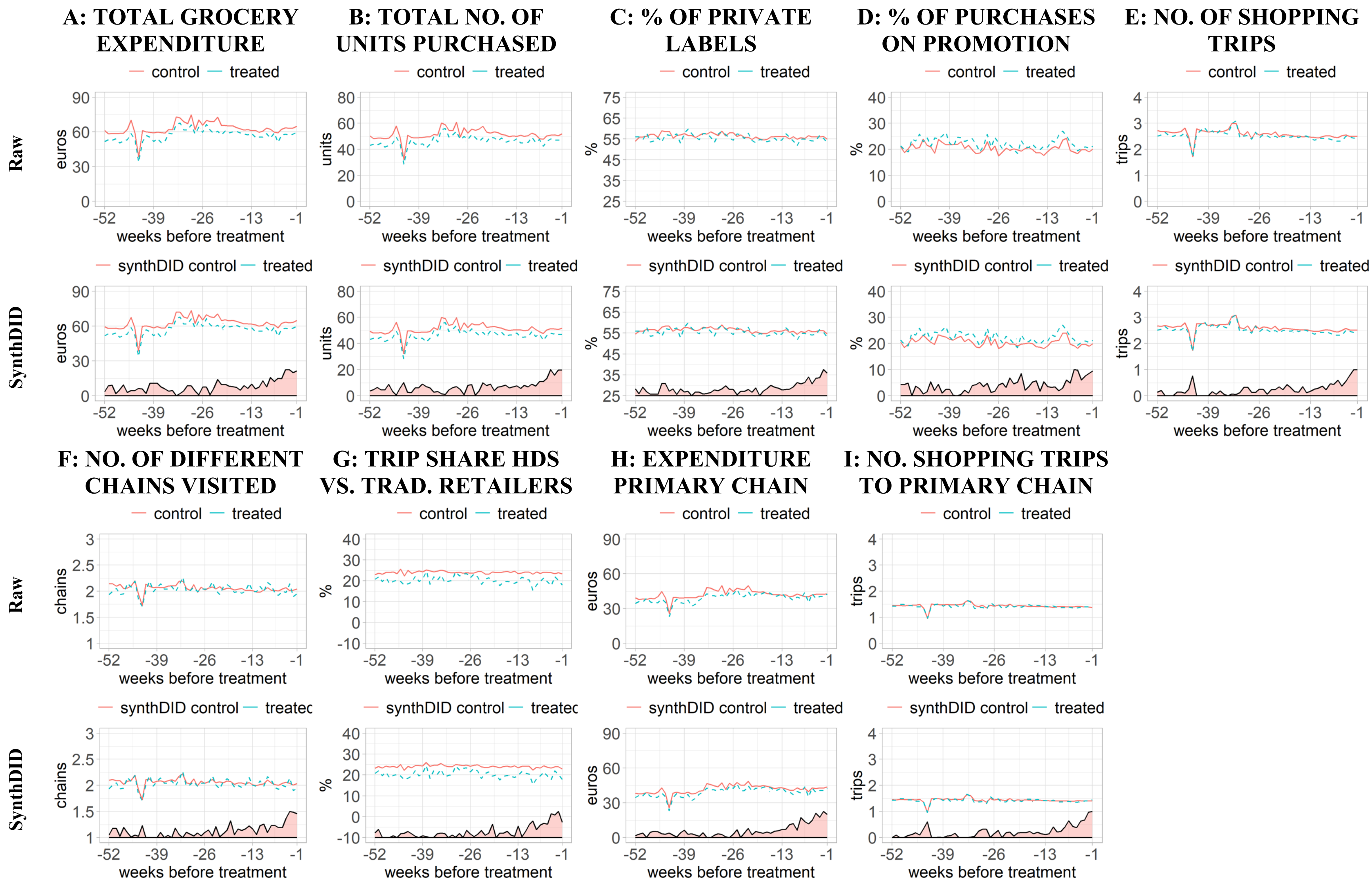
Notes: Dashed (solid) lines represent the weekly average for treated (control) households. The weighted average for the SynthDID control households is weighted as described in the "Synthetic Difference-in-Differences" section. The time weights used to average pre-treatment periods in the SynthDID approach are at the bottom of each plot.

FIGURE WA5: COHORT 4



Notes: Dashed (solid) lines represent the weekly average for treated (control) households. The weighted average for the SynthDID control households is weighted as described in the "Synthetic Difference-in-Differences" section. The time weights used to average pre-treatment periods in the SynthDID approach are at the bottom of each plot.

FIGURE WA6: COHORT 5



Notes: Dashed (solid) lines represent the weekly average for treated (control) households. The weighted average for the SynthDID control households is weighted as described in the "Synthetic Difference-in-Differences" section. The time weights used to average pre-treatment periods in the SynthDID approach are at the bottom of each plot.

WEB APPENDIX D: INTERVAL ESTIMATE CURVES

Figure WA7 shows interval estimates curves for each outcome variable following recent suggestions by McShane et al. (2024). The grey area denotes the confidence interval (interval estimate) for all levels from 0% to 99.9% (right y-axis). The peak (0%) denotes the point estimate (e.g., .40 for total grocery expenditure) and the bounds of a traditional 95% confidence interval (CI) are shown towards the bottom.

FIGURE WA7: INTERVAL ESTIMATE CURVES

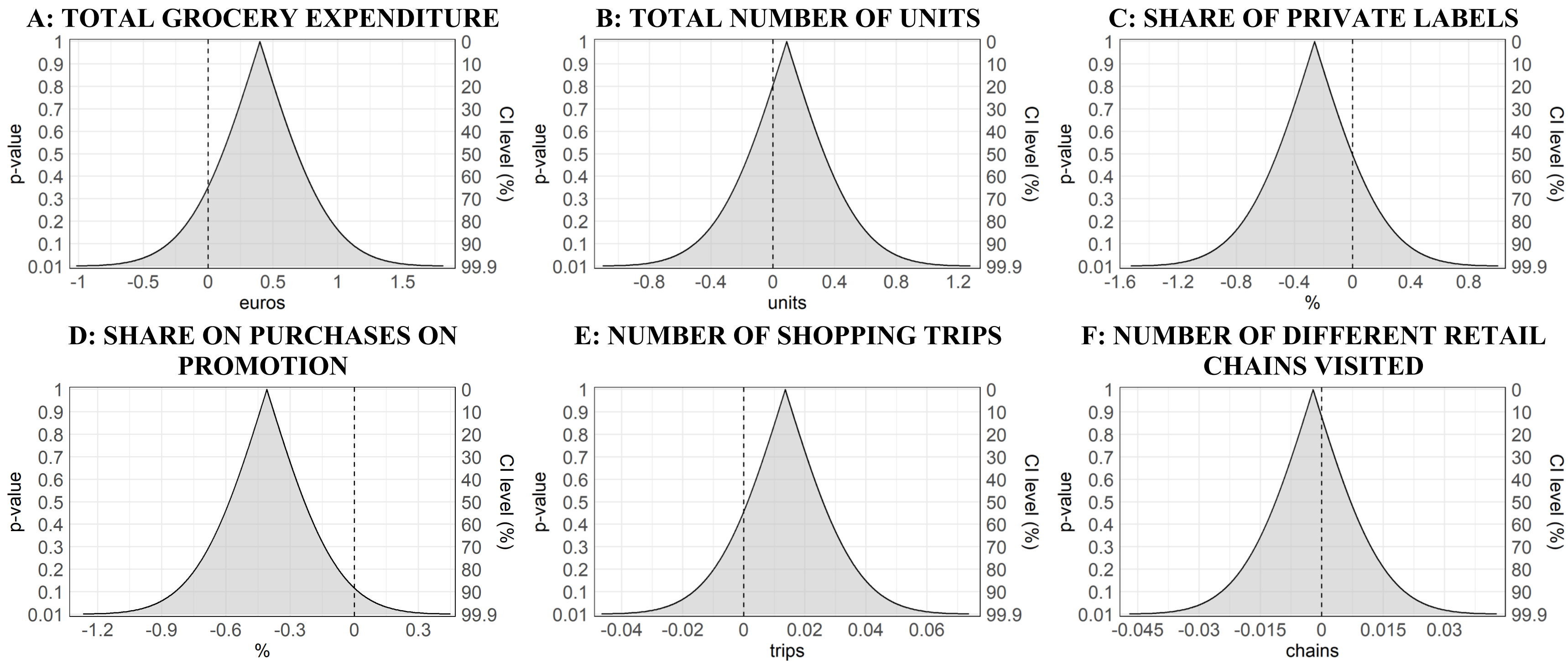
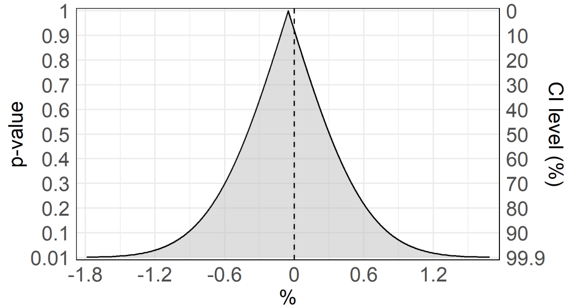
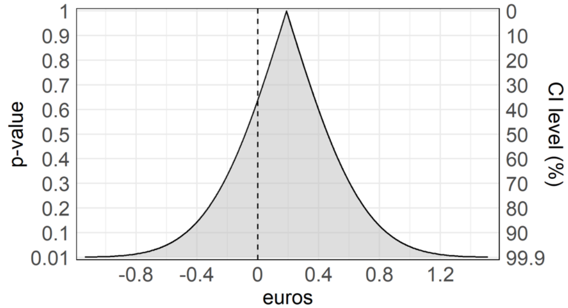


FIGURE WA7 (CONTINUED)

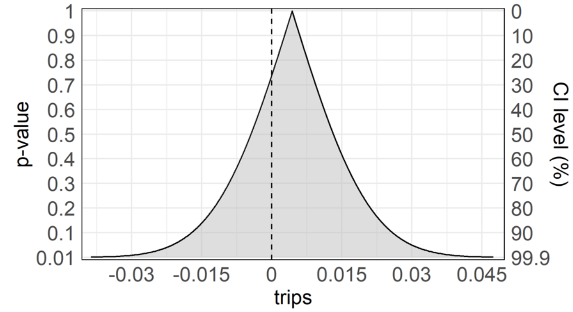
G: TRIP SHARE HDS VS. TRADITIONAL RETAILERS



H: GROCERY EXPENDITURE AT PRIMARY CHAIN



I: NUMBER OF SHOPPING TRIPS TO PRIMARY CHAIN



Notes: The curve plots the confidence interval (CI) for each dependent variable for the level indicated by the right y-axis. Similarly, it plots the p-value for each variable against the target hypothesis of the value on the x-axis.

WEB APPENDIX E: SYNTHETIC CONTROL METHOD

As discussed in the main text, an alternative analysis method would be to use synthetic control (SC). SC can be seen as a middle ground between DID and SynthDID in terms of the number of weights and parameters that are estimated. However, and crucially, the lack of regularization in terms of control weights vis-à-vis SynthDID may lead to overfitting. This issue seems particularly relevant in our setting as it has many controls. Indeed, as Abadie (2021) and Li and Sonnier (2023) write, the risk with the SC method in combination with a large donor pool is that of overfitting. In particular, a lack of regularization means that the SC method tries to find controls that predict the pre-treatment DV without penalizing the control weights. In practice, this could lead to an over-reliance on a small subset of controls, rather than more nuanced down/upweighting of the SynthDID approach (see also Arkhangelsky et al. 2021, p. 4094).

To understand this further, we compare our results with that of SC and compute several complementary statistics to assess overfitting. Table WA3 shows the results of SC and contrasts it with the results of our focal SynthDID model and DID. SC yields substantially larger parameter estimates than our focal estimates and some confidence intervals no longer include zero. Yet, these effects may be caused by overfitting. As a first test to understand this, we calculate the prediction mean squared error (PMSE) suggested by Li and Van den Bulte (2023), by splitting the pre-treatment period and using the last 26 weeks of the pre-treatment period as pseudo post-treatment period. In our setting the PMSE is substantially higher for SC than the focal SynthDID ($PMSE_{focal}$ / $PMSE_{SC}$ = .16 on average across cohorts and dependent variables). Next, to formally assess over-reliance on a small subset of controls, we calculate several statistics based on the control weights obtained in our analysis. In Table WA4, we contrast the different methods in terms of their approach (column 1) and several descriptives derived from the realized values for the control weights in our setting (columns 2-4). As shown in column 2, we find that SC relies on very few controls (3.27%) and discards roughly 96.73% of the control observations. Put differently, with SC, few controls uniquely determine the behavior of the control group post-treatment. For SynthDID, we observe that 97.94% of the controls receive a positive weight, such that only few controls (2.06%) are completely discarded. In column 3, we calculate the ratio of the weight for each non-zero weighted control over the DID weight (1/N). For SC, we see that the used (non-zero-weight) controls on average are 37.91 times more influential than they would have been in a DID model (vis-à-vis 1.023 for SynthDID vs. DID). Last, in column 4, we also calculate a measure of dispersion of the control weights. More specifically, we calculate the Herfindahl-Hirschman index (HHI) per method and normalize it such that the HHI for DID equals 1. We find a value of 335.32 for SC, whereas SynthDID stays relatively close to DID, with a value of 1.12. In summary, in our setting, the control weights in SynthDID are distributed relatively evenly across the large majority of controls, whereas with SC, few controls receive relatively large weights, creating an over-reliance on these weights.

Last, we test the stability of the estimates by contrasting the estimates obtained with SC with those of random samples of a smaller size (that are less prone to overfitting). In particular, Table WA5 shows the statistics on the estimates from 500 random control group samples, where we each time restrict the number of potential control units to be the same as the number of treated units. The means of the parameter estimates coming from the 500 samples (in Panel A) are in line with DID and SynthDID. Finally, we also observe a much higher standard deviation in effect size for SC (vis-à-vis DID and SynthDID) across the 500 samples (see Panel B), suggesting that the (potential over-)reliance on few controls influences the results substantially.

TABLE WA3: SC ESTIMATES VS. FOCAL AND DID ESTIMATES

	Focal	DID	SC
Total grocery expenditure (in euros)	.40 [-.45, 1.24]	.70 [-.20, 1.59]	.58 [-.58, 2.07]
Total number of units purchased	.09 [-.62, .80]	.39 [-.37, 1.15]	**1.39 [.38, 2.51]**
Share of private labels (%)	-.26 [-1.02, .49]	-.28 [-1.05, .48]	-.85 [-1.38, .17]
Share of purchases on promotion (%)	-.41 [-.92, .10]	-.42 [-.94, .09]	**-.40 [-1.15, -.12]**
Number of shopping trips	.01 [-.02, .05]	.02 [-.02, .06]	**-.10 [-.15, -.03]**
Number of different retail chains visited	.00 [-.03, .02]	.00 [-.03, .03]	**-.06 [-.09, -.04]**
Trip share HDs vs. traditional retailers (%)	-.05 [-1.09, .98]	-.22 [-1.26, .82]	-.30 [-.67, .08]
Grocery expenditure at primary chain (in euros)	.19 [-.60, .98]	.55 [-.30, 1.41]	**1.81 [.52, 3.85]**
Number of shopping trips to primary chain	.00 [-.02, .03]	.01 [-.02, .04]	-.03 [-.07, .01]

Notes: 95% confidence interval in brackets; confidence intervals excluding zero in bold. For SC, the Jackknife procedure is discouraged (see Arkhangelsky et al. 2021) and we use the subsampling method as discussed in Li (2020). We set the subsample size to 26 weeks (i.e., half the number of pre-treatment periods) and use 10,000 subsample simulations. HD = hard discounter, coh. = cohort.

TABLE WA4: SUMMARY STATISTICS CONTROL WEIGHTS FOR SC VS. SYNTHDID AND DID

	Approach regarding control weights	% of controls with non-zero weight	Average weight of non-zero controls divided by the average weight in DID	HHI (normalized by HHI of DID)
DID	$1/N_c$	100%	1	1
SC	Estimated	3.27%	37.91	335.32
SynthDID	Estimated with regularization	97.94%	1.023	1.12

Notes: Values displayed are averages across all nine DVs and five cohorts.

TABLE WA5: STABILITY OF ESTIMATES

A: ESTIMATES MAIN ANALYSES VS. MEAN OF 500 RANDOM SAMPLES

	SynthDID		DID		SC	
	Main	**Samples**	**Main**	**Samples**	**Main**	**Samples**
Total grocery expenditure (in euros)	.40	.42	.70	.71	.58	.02
Total number of units purchased	.09	-.01	.39	.40	1.39	-.34
Share of private labels (%)	-.26	-.20	-.28	-.28	-.85	-.31
Share of purchases on promotion (%)	-.41	-.55	-.42	-.42	-.40	-.25
Number of shopping trips	.01	.01	.02	.02	-.10	-.00
Number of different retail chains visited	.00	-.01	.00	.00	-.06	-.01
Trip share HDs vs. traditional retailers (%)	-.05	-.27	-.22	-.22	-.30	-.28
Grocery expenditure at primary chain (in euros)	.19	.16	.55	.56	1.81	.14
Number of shopping trips to primary chain	.00	-.00	.01	.01	-.03	.01

B: STANDARD DEVIATION OF 500 RANDOM SAMPLES

	SynthDID	**DID**	**SC**
Total grocery expenditure (in euros)	.39	.37	.81
Total number of units purchased	.34	.33	.66
Share of private labels (%)	.38	.32	.48
Share of purchases on promotion (%)	.28	.23	.40
Number of shopping trips	.02	.02	.03
Number of different retail chains visited	.02	.01	.02
Trip share HDs vs. traditional retailers (%)	.40	.40	.52
Grocery expenditure at primary chain (in euros)	.39	.38	.77
Number of shopping trips to primary chain	.01	.01	.03

Notes: HD = hard discounter. The estimates in the columns "Main" are the results of our main analyses, as reported in the main text/this Web Appendix.

WEB APPENDIX F: POWER TO DETECT MEANINGFUL EFFECTS

The estimates in Table 7 in the main text are small and statistically indistinguishable from zero. However, null findings may be driven by the lack of an effect or insufficient statistical power to detect economically meaningful effects. There are two noteworthy characteristics of our setting that influence the calculation of MDEs. First, we employ panel data, therefore we must use a MDE calculation that is robust to serial correlation. Second, the treatment is implemented in a staggered manner, requiring us to use an MDE calculation that is robust to this staggered implementation. We note that while the literature has developed MDEs using DID designs in a panel setting (Burlig, Preonas, and Woerman 2020), with extensions to staggered treatment (Schochet 2022), generalizations that incorporate SynthDID do not exist to the best of our knowledge. We consider developing power calculations for SynthDID with staggered adoption to be beyond the scope of this paper and, supported by the largely identical results obtained with the SynthDID and DID methods, report the MDEs based on a staggered two-way fixed effects DID specification (as used in the robustness check in the main text).

To calculate the MDE, we follow the procedure outlined in Equation 12 by Schochet (2022, p. 382) to calculate the variance of the staggered treatment effect defined by:

$$\mathrm{Var}\left(\hat{\tau}^{\mathrm{DID}}\right) = \frac{1}{\left(\sum_{r=1}^{R} A_r\right)^2} \sum_{r=1}^{R} A_r^2 \, \mathrm{Var}\left(\hat{\tau}_r^{\mathrm{DID}}\right), \tag{W1}$$

where A_r equals the weight of cohort r, determined by the respective number of treated households in cohort r. To determine $\mathrm{Var}\left(\hat{\tau}_r^{\mathrm{DID}}\right)$, we Equation 2 of Burlig, Preonas, and Woerman (2020, p. 3):

$$\mathrm{Var}\left(\hat{\tau}_r^{\mathrm{DID}}\right) = \frac{1}{P(1-P)J}\left[\left(\frac{\ell_{\min} + \ell_{\max}}{\ell_{\min} * \ell_{\max}}\right)\sigma_w^2 + \left(\frac{\ell_{\min} - 1}{\ell_{\min}}\right)\psi^B + \left(\frac{\ell_{\max} - 1}{\ell_{\max}}\right)\psi^A - 2\psi^X\right], \tag{W2}$$

where P is the proportion of observations of households in treatment municipalities after the ban is active; J is the total number of households; $\ell_{\min}$ and $\ell_{\max}$ refer to the number of weeks pre- and post-treatment, respectively; σ_w^2 indicates the variance of the error term in a two-way fixed effects DID framework; and ψ^B, ψ^A, and ψ^X are the average pre-treatment, post-treatment, and across-period covariance between residuals at t and t+1 of the given time period (of the same household), respectively. As before, we set $\ell_{\min}$ and $\ell_{\max}$ to 52 and derive all other quantities from the data.[1]

Then, as a last step, we calculate the MDE for each dependent variable as:

$$\mathrm{MDE} = \mathrm{M}_{\mathrm{J}-2}\sqrt{\mathrm{Var}(\hat{\tau}^{\mathrm{DID}})}, \tag{W3}$$

where $\mathrm{M}_{\mathrm{J}-2}$ is the sum of the t-values for alpha (significance level) and beta (power). We set alpha (significance) and beta (power) to conventional values of .05 and .80, resulting in an $\mathrm{M}_{\mathrm{J}-2}$ of 2.8 ($\mathrm{t}^{\mathrm{J}}_{\alpha/2} + \mathrm{t}^{\mathrm{J}}_{1-\beta}$ = 1.96 + .84 = 2.8). We present the MDEs in Table WA6.

As Table WA6 indicates, the MDEs for the dependent variables are relatively small. These results lend credibility to our estimates, which are smaller than the MDEs. Moreover, for most treatment effects, the MDEs lie within the confidence intervals. That is, the MDEs effectively tighten the bounds of our estimates (cf. the confidence interval itself) because we should be able to find effects that are as large as the MDE. We thus gain meaningful insights into the upper bound of the treatment effect. In conclusion, both the estimated treatment effect and the MDE boundaries are small in magnitude, confirming the lack of a sizable effect in our setting.

[1] To derive σ^2_{w}, ψ^{B}, ψ^{A}, and ψ^{X} , we retain the residuals of the estimated two-way fixed effects models. We then, (1) calculate the variance of these residuals to obtain σ^2_{w}, (2) calculate the within-unit covariance of the pre-period (ψ^{B}), post-period (ψ^{A}), and across period (ψ^{X}) residuals.

TABLE WA6: MINIMUM DETECTABLE EFFECTS (MDEs)

	MDE	DID estimate
Total grocery expenditure (in euros)	1.09	.70 [-.20, 1.59]
Total number of units purchased	.94	.39 [-.37, 1.15]
Share of private labels (%)	1.03	-.28 [-1.05, .48]
Share of purchases on promotion (%)	.67	-.42 [-.94, .09]
Number of shopping trips	.05	.02 [-.02, .06]
Number of different retail chains visited	.04	.00 [-.03, .03]
Trip share HDs vs. traditional retailers (%)	1.38	-.22 [-1.26, .82]
Grocery expenditure at primary chain	1.17	.55 [-.30, 1.41]
Number of shopping trips to primary chain	.04	.01 [-.02, .04]

Notes: Notes: HD = hard discounter. MDEs are presented in absolute terms. 95% confidence interval of the DID estimates in brackets. DID estimates and confidence intervals come from Table 8 in the main text.

WEB APPENDIX G: IMPUTATION OF OPT-IN CHOICE

As discussed in the “Setting and Data” section of the main text, we obtain data on households’ opt-in choices (i.e., use of a mailbox sticker) from GfK and NOM, a Dutch media research organization, which survey a sample of GfK panel households every year. In the first week of 2021, just after the introduction of the latest ban, we ran a separate survey among all treated households that were still active in the panel, as well as a random sample of control households to supplement the GfK/NOM data. The survey asked households if they used a mailbox sticker and, if so, when they started using it. Due to panel attrition and nonresponse, the data on the choice to opt in were available for 60% of the households in the treatment municipalities. We impute the opt-in choice for the remaining 40% of households using a random forest algorithm with the missForest package in R (Stekhoven and Buehlmann 2012). As inputs, we use (1) households’ opt-out choice in the period before treatment,[2] (2) the opt-in choice of the geographically nearest three panel households, (3) household demographics obtained from GfK (household size, age of head of household, income class, and social class), (4) neighborhood characteristics obtained from Statistics Netherlands (percentage of households on social welfare, percentage of Western immigrants, percentage of non-Western immigrants, home ownership rate, average home value, population density, distance to the nearest grocery store, number of grocery stores in a 3-km radius, distance to the nearest department store, and number of department stores in a 5-km radius), and (5) municipality dummy variables. We obtain a hit rate of 94.8%.

[2] For households for which the data do not include their opt-out choice in the period before treatment either, we impute their opt-out choice using a similar approach.

WEB APPENDIX H: COMPOSITION OF SUBGROUPS

Table WA7 shows descriptive statistics of the household demographics of the three subgroups across all cohorts. Tables WA8-WA12 show similar statistics per cohort.

TABLE WA7: DEMOGRAPHICS TREATED HOUSEHOLDS PER SUBGROUP, ACROSS COHORTS

A: COMPLIERS (HOUSEHOLDS WITHOUT A STICKER) (N = 344)

	Mean	SD	Min	Max
Household size[a]	2.15	1.17	1	5
Age of head of household[b]	7.74	2.15	2	11
Income class[c]	10.04	5.22	1	19
Social class[d]	3.30	1.34	1	5

B: DEFIERS (HOUSEHOLDS THAT OPT IN) (N = 372)

	Mean	SD	Min	Max
Household size[a]	2.23	1.17	1	5
Age of head of household[b]	8.40	1.93	1	11
Income class[c]	9.59	4.98	1	19
Social class[d]	2.81	1.34	1	5

C: ALWAYS-TAKERS (HOUSEHOLDS THAT ALREADY OPTED OUT) (N = 176)

	Mean	SD	Min	Max
Household size[a]	1.69	1.01	1	5
Age of head of household[b]	7.09	2.48	2	11
Income class[c]	9.28	4.97	1	19
Social class[d]	3.65	1.26	1	5

[a] 1 = 1 household member, 2 = 2 household members, 3 = 3 household members, 4 = 4 household members, 5 = 5 or more household members.

[b] 1 = 12-19 y.o. (years old), 2 = 20-24 y.o., 3 = 25-29 y.o., 4 = 30-34 y.o., 5 = 35-39 y.o., 6 = 40-44 y.o., 7 = 45-49 y.o., 8 = 50-54 y.o., 9 = 55-64 y.o., 10 = 65-74 y.o., 11 = 75 y.o. or older.

[c] Net income per month: 1 = below 700, 2 = 700-900, 3 = 900-1,100, 4 = 1,100-1,300, 5 = 1,300-1,500, 6 = 1,500-1,700, 7 = 1,700-1,900, 8 = 1,900-2,100, 9 = 2,100-2,300, 10 = 2,300-2,500, 11 = 2,500-2,700, 12 = 2,700-2,900, 13 = 2,900-3,100, 14 = 3,100-3,300, 15 = 3,300-3,500, 16 = 3,500-3,700, 17 = 3,700-3,900, 18 = 3,900-4,100, 19 = 4,100 or more.

[d] 1 = D (lower), 2 = C, 3 = B-minus, 4 = B-plus, 5 = A (upper); based on the education level and occupation of the head of the household.

TABLE WA8: DEMOGRAPHICS TREATED HOUSEHOLDS PER SUBGROUP, COHORT 1

A: COMPLIERS (HOUSEHOLDS WITHOUT A STICKER) (N = 71)

	Mean	SD	Min	Max
Household size[a]	1.73	0.97	1	5
Age of head of household[b]	8.42	1.93	2	11
Income class[c]	8.66	4.94	2	19
Social class[d]	2.87	1.41	1	5

B: DEFIERS (HOUSEHOLDS THAT OPT IN) (N = 79)

	Mean	SD	Min	Max
Household size[a]	1.92	1.07	1	5
Age of head of household[b]	9.13	1.46	5	11
Income class[c]	8.51	5.04	2	19
Social class[d]	2.66	1.40	1	5

C: ALWAYS-TAKERS (HOUSEHOLDS THAT ALREADY OPTED OUT) (N = 24)

	Mean	SD	Min	Max
Household size[a]	1.92	1.14	1	5
Age of head of household[b]	7.46	2.30	2	11
Income class[c]	10.79	5.71	1	19
Social class[d]	4.21	0.98	2	5

[a] 1 = 1 household member, 2 = 2 household members, 3 = 3 household members, 4 = 4 household members, 5 = 5 or more household members.

[b] 1 = 12-19 y.o. (years old), 2 = 20-24 y.o., 3 = 25-29 y.o., 4 = 30-34 y.o., 5 = 35-39 y.o., 6 = 40-44 y.o., 7 = 45-49 y.o., 8 = 50-54 y.o., 9 = 55-64 y.o., 10 = 65-74 y.o., 11 = 75 y.o. or older.

[c] Net income per month: 1 = below 700, 2 = 700-900, 3 = 900-1,100, 4 = 1,100-1,300, 5 = 1,300-1,500, 6 = 1,500-1,700, 7 = 1,700-1,900, 8 = 1,900-2,100, 9 = 2,100-2,300, 10 = 2,300-2,500, 11 = 2,500-2,700, 12 = 2,700-2,900, 13 = 2,900-3,100, 14 = 3,100-3,300, 15 = 3,300-3,500, 16 = 3,500-3,700, 17 = 3,700-3,900, 18 = 3,900-4,100, 19 = 4,100 or more.

[d] 1 = D (lower), 2 = C, 3 = B-minus, 4 = B-plus, 5 = A (upper); based on the education level and occupation of the head of the household.

TABLE WA9: DEMOGRAPHICS TREATED HOUSEHOLDS PER SUBGROUP, COHORT 2

A: COMPLIERS (HOUSEHOLDS WITHOUT A STICKER) (N = 92)

	Mean	SD	Min	Max
Household size[a]	2.32	1.17	1	5
Age of head of household[b]	7.63	2.38	3	11
Income class[c]	11.03	5.38	1	19
Social class[d]	3.34	1.28	1	5

B: DEFIERS (HOUSEHOLDS THAT OPT IN) (N = 135)

	Mean	SD	Min	Max
Household size[a]	2.41	1.25	1	5
Age of head of household[b]	8.01	1.99	1	11
Income class[c]	9.69	5.04	1	19
Social class[d]	2.93	1.28	1	5

C: ALWAYS-TAKERS (HOUSEHOLDS THAT ALREADY OPTED OUT) (N = 66)

	Mean	SD	Min	Max
Household size[a]	1.71	1.05	1	5
Age of head of household[b]	6.83	2.64	2	11
Income class[c]	9.95	4.78	3	19
Social class[d]	3.92	1.18	1	5

[a] 1 = 1 household member, 2 = 2 household members, 3 = 3 household members, 4 = 4 household members, 5 = 5 or more household members.

[b] 1 = 12-19 y.o. (years old), 2 = 20-24 y.o., 3 = 25-29 y.o., 4 = 30-34 y.o., 5 = 35-39 y.o., 6 = 40-44 y.o., 7 = 45-49 y.o., 8 = 50-54 y.o., 9 = 55-64 y.o., 10 = 65-74 y.o., 11 = 75 y.o. or older.

[c] Net income per month: 1 = below 700, 2 = 700-900, 3 = 900-1,100, 4 = 1,100-1,300, 5 = 1,300-1,500, 6 = 1,500-1,700, 7 = 1,700-1,900, 8 = 1,900-2,100, 9 = 2,100-2,300, 10 = 2,300-2,500, 11 = 2,500-2,700, 12 = 2,700-2,900, 13 = 2,900-3,100, 14 = 3,100-3,300, 15 = 3,300-3,500, 16 = 3,500-3,700, 17 = 3,700-3,900, 18 = 3,900-4,100, 19 = 4,100 or more.

[d] 1 = D (lower), 2 = C, 3 = B-minus, 4 = B-plus, 5 = A (upper); based on the education level and occupation of the head of the household.

TABLE WA10: DEMOGRAPHICS TREATED HOUSEHOLDS PER SUBGROUP, COHORT 3

A: COMPLIERS (HOUSEHOLDS WITHOUT A STICKER) (N = 74)

	Mean	SD	Min	Max
Household size[a]	2.35	1.16	1	5
Age of head of household[b]	7.72	2.01	3	11
Income class[c]	10.65	5.51	3	19
Social class[d]	3.41	1.37	1	5

B: DEFIERS (HOUSEHOLDS THAT OPT IN) (N = 65)

	Mean	SD	Min	Max
Household size[a]	2.17	1.13	1	5
Age of head of household[b]	8.26	1.92	3	11
Income class[c]	9.97	4.87	1	19
Social class[d]	2.98	1.30	1	5

C: ALWAYS-TAKERS (HOUSEHOLDS THAT ALREADY OPTED OUT) (N = 44)

	Mean	SD	Min	Max
Household size[a]	1.68	0.96	1	5
Age of head of household[b]	7.16	2.36	3	11
Income class[c]	8.75	5.45	2	19
Social class[d]	3.34	1.27	1	5

[a] 1 = 1 household member, 2 = 2 household members, 3 = 3 household members, 4 = 4 household members, 5 = 5 or more household members.

[b] 1 = 12-19 y.o. (years old), 2 = 20-24 y.o., 3 = 25-29 y.o., 4 = 30-34 y.o., 5 = 35-39 y.o., 6 = 40-44 y.o., 7 = 45-49 y.o., 8 = 50-54 y.o., 9 = 55-64 y.o., 10 = 65-74 y.o., 11 = 75 y.o. or older.

[c] Net income per month: 1 = below 700, 2 = 700-900, 3 = 900-1,100, 4 = 1,100-1,300, 5 = 1,300-1,500, 6 = 1,500-1,700, 7 = 1,700-1,900, 8 = 1,900-2,100, 9 = 2,100-2,300, 10 = 2,300-2,500, 11 = 2,500-2,700, 12 = 2,700-2,900, 13 = 2,900-3,100, 14 = 3,100-3,300, 15 = 3,300-3,500, 16 = 3,500-3,700, 17 = 3,700-3,900, 18 = 3,900-4,100, 19 = 4,100 or more.

[d] 1 = D (lower), 2 = C, 3 = B-minus, 4 = B-plus, 5 = A (upper); based on the education level and occupation of the head of the household.

TABLE WA11: DEMOGRAPHICS TREATED HOUSEHOLDS PER SUBGROUP, COHORT 4

A: COMPLIERS (HOUSEHOLDS WITHOUT A STICKER) (N = 19)

	Mean	SD	Min	Max
Household size[a]	2.37	1.34	1	5
Age of head of household[b]	7.58	2.14	4	11
Income class[c]	8.63	3.95	2	19
Social class[d]	3.21	1.27	1	5

B: DEFIERS (HOUSEHOLDS THAT OPT IN) (N = 6)

	Mean	SD	Min	Max
Household size[a]	2.00	1.10	1	4
Age of head of household[b]	8.67	1.86	5	10
Income class[c]	9.50	6.06	4	19
Social class[d]	3.00	1.67	1	5

C: ALWAYS-TAKERS (HOUSEHOLDS THAT ALREADY OPTED OUT) (N = 4)

	Mean	SD	Min	Max
Household size[a]	1.25	0.50	1	2
Age of head of household[b]	4.75	1.50	3	6
Income class[c]	7.00	5.94	2	15
Social class[d]	3.75	0.96	3	5

[a] 1 = 1 household member, 2 = 2 household members, 3 = 3 household members, 4 = 4 household members, 5 = 5 or more household members.

[b] 1 = 12-19 y.o. (years old), 2 = 20-24 y.o., 3 = 25-29 y.o., 4 = 30-34 y.o., 5 = 35-39 y.o., 6 = 40-44 y.o., 7 = 45-49 y.o., 8 = 50-54 y.o., 9 = 55-64 y.o., 10 = 65-74 y.o., 11 = 75 y.o. or older.

[c] Net income per month: 1 = below 700, 2 = 700-900, 3 = 900-1,100, 4 = 1,100-1,300, 5 = 1,300-1,500, 6 = 1,500-1,700, 7 = 1,700-1,900, 8 = 1,900-2,100, 9 = 2,100-2,300, 10 = 2,300-2,500, 11 = 2,500-2,700, 12 = 2,700-2,900, 13 = 2,900-3,100, 14 = 3,100-3,300, 15 = 3,300-3,500, 16 = 3,500-3,700, 17 = 3,700-3,900, 18 = 3,900-4,100, 19 = 4,100 or more.

[d] 1 = D (lower), 2 = C, 3 = B-minus, 4 = B-plus, 5 = A (upper); based on the education level and occupation of the head of the household.

TABLE WA12: DEMOGRAPHICS TREATED HOUSEHOLDS PER SUBGROUP, COHORT 5

A: COMPLIERS (HOUSEHOLDS WITHOUT A STICKER) (N = 88)

	Mean	SD	Min	Max
Household size[a]	2.10	1.23	1	5
Age of head of household[b]	7.36	2.10	3	11
Income class[c]	9.91	5.04	2	19
Social class[d]	3.53	1.29	1	5

B: DEFIERS (HOUSEHOLDS THAT OPT IN) (N = 87)

	Mean	SD	Min	Max
Household size[a]	2.29	1.10	1	5
Age of head of household[b]	8.41	2.05	3	11
Income class[c]	10.16	4.80	1	19
Social class[d]	2.63	1.35	1	5

C: ALWAYS-TAKERS (HOUSEHOLDS THAT ALREADY OPTED OUT) (N = 38)

	Mean	SD	Min	Max
Household size[a]	1.55	0.95	1	5
Age of head of household[b]	7.45	2.45	2	11
Income class[c]	8.00	3.75	3	19
Social class[d]	3.18	1.35	1	5

[a] 1 = 1 household member, 2 = 2 household members, 3 = 3 household members, 4 = 4 household members, 5 = 5 or more household members.

[b] 1 = 12-19 y.o. (years old), 2 = 20-24 y.o., 3 = 25-29 y.o., 4 = 30-34 y.o., 5 = 35-39 y.o., 6 = 40-44 y.o., 7 = 45-49 y.o., 8 = 50-54 y.o., 9 = 55-64 y.o., 10 = 65-74 y.o., 11 = 75 y.o. or older.

[c] Net income per month: 1 = below 700, 2 = 700-900, 3 = 900-1,100, 4 = 1,100-1,300, 5 = 1,300-1,500, 6 = 1,500-1,700, 7 = 1,700-1,900, 8 = 1,900-2,100, 9 = 2,100-2,300, 10 = 2,300-2,500, 11 = 2,500-2,700, 12 = 2,700-2,900, 13 = 2,900-3,100, 14 = 3,100-3,300, 15 = 3,300-3,500, 16 = 3,500-3,700, 17 = 3,700-3,900, 18 = 3,900-4,100, 19 = 4,100 or more.

[d] 1 = D (lower), 2 = C, 3 = B-minus, 4 = B-plus, 5 = A (upper); based on the education level and occupation of the head of the household.

WEB APPENDIX I: RESULTS OF SUBGROUP ANALYSES

TABLE WA13: SYNTHDID ESTIMATES OF THE IMPACT OF INTRODUCING A BAN

A: COMPLIERS (HOUSEHOLDS WITHOUT STICKER)

	Estimate		95% Confidence Interval		
	$\hat{\tau}$	Jackknife SE	Lower Bound	Upper Bound	*p*-Value
Total grocery expenditure (in euros)	.18	.66	-1.11	1.46	.79
Total number of units purchased	-.08	.52	-1.10	.95	.88
Share of private labels (%)	-.55	.69	-1.91	.81	.43
Share of purchases on promotion (%)	-.97	.44	-1.82	-.11	.03
Number of shopping trips	.01	.03	-.05	.07	.72
Number of different retail chains visited	-.02	.02	-.07	.02	.30
Trip share HDs vs. traditional retailers (%)	-1.04	1.03	-3.06	.97	.31
Grocery expenditure at primary chain	-.02	.64	-1.28	1.23	.97
Number of shopping trips to primary chain	.00	.02	-.04	.04	.87

Notes: HD = hard discounter. Number of treated (control) households: 344 (32,303) for variables unconditional on a trip observed in a week, and 136 (13,474) for variables conditional on a trip observed in a week.

B: DEFIERS (HOUSEHOLDS THAT OPT IN)

	Estimate		95% Confidence Interval		
	$\hat{\tau}$	Jackknife SE	Lower Bound	Upper Bound	*p*-Value
Total grocery expenditure (in euros)	.70	.72	-.70	2.10	.33
Total number of units purchased	-.03	.63	-1.28	1.21	.96
Share of private labels (%)	-.03	.47	-.94	.88	.94
Share of purchases on promotion (%)	.27	.37	-.45	.99	.46
Number of shopping trips	.00	.03	-.06	.05	.96
Number of different retail chains visited	.01	.02	-.03	.05	.57
Trip share HDs vs. traditional retailers (%)	.21	.42	-.61	1.04	.61
Grocery expenditure at primary chain	.79	.62	-.43	2.01	.21
Number of shopping trips to primary chain	.01	.02	-.02	.05	.49

Notes: HD = hard discounter. Number of treated (control) households: 372 (32,303) for variables unconditional on a trip observed in a week, and 168 (13,474) for variables conditional on a trip observed in a week.

TABLE WA13 (CONTINUED)
C: ALWAYS-TAKERS (HOUSEHOLDS THAT ALREADY OPTED OUT)

	Estimate		95% Confidence Interval		
	$\hat{\tau}$	Jackknife SE	Lower Bound	Upper Bound	*p*-Value
Total grocery expenditure (in euros)	.19	.87	-1.52	1.90	.83
Total number of units purchased	.67	.69	-.67	2.01	.33
Share of private labels (%)	-.25	1.12	-2.45	1.94	.82
Share of purchases on promotion (%)	-1.07	.64	-2.33	.18	.09
Number of shopping trips	.05	.04	-.03	.13	.20
Number of different retail chains visited	.01	.04	-.06	.09	.75
Trip share HDs vs. traditional retailers (%)	1.53	1.88	-2.16	5.21	.42
Grocery expenditure at primary chain	-.66	.89	-2.40	1.07	.45
Number of shopping trips to primary chain	.00	.03	-.06	.06	.98

Notes: HD = hard discounter. Number of treated (control) households: 176 (32,303) for variables unconditional on a trip observed in a week, and 57 (13,474) for variables conditional on a trip observed in a week.

D: NON-COMPLIERS (HOUSEHOLDS THAT OPT IN AND HOUSEHOLDS THAT ALREADY OPTED OUT COMBINED)

	Estimate		95% Confidence Interval		
	$\hat{\tau}$	Jackknife SE	Lower Bound	Upper Bound	*p*-Value
Total grocery expenditure (in euros)	.54	.56	-.57	1.64	.34
Total number of units purchased	.19	.49	-.76	1.14	.69
Share of private labels (%)	-.09	.45	-.97	.79	.84
Share of purchases on promotion (%)	-.07	.32	-.70	.56	.83
Number of shopping trips	.02	.02	-.03	.06	.51
Number of different retail chains visited	.01	.02	-.02	.04	.52
Trip share HDs vs. traditional retailers (%)	.55	.57	-.57	1.67	.34
Grocery expenditure at primary chain	.32	.51	-.68	1.33	.53
Number of shopping trips to primary chain	.01	.02	-.02	.04	.58

Notes: HD = hard discounter. Number of treated (control) households: 548 (32,303) for variables unconditional on a trip observed in a week, and 225 (13,474) for variables conditional on a trip observed in a week.

WEB APPENDIX J: TREATMENT HETEROGENEITY

TABLE WA14: SYNTHDID ESTIMATES OF THE IMPACT OF INTRODUCING A BAN ACROSS INCOME GROUPS

A: HOUSEHOLDS WITH LOW INCOME

	Estimate		95% Confidence Interval		
	$\hat{\tau}$	Jackknife SE	Lower Bound	Upper Bound	*p*-Value
Total grocery expenditure (in euros)	-.12	.51	-1.11	.87	.81
Total number of units purchased	-.38	.38	-1.13	.37	.32
Share of private labels (%)	.52	.54	-.54	1.58	.33
Share of purchases on promotion (%)	-.17	.41	-.98	.64	.68
Number of shopping trips	.03	.03	-.03	.08	.34
Number of different retail chains visited	.01	.02	-.03	.05	.64
Trip share HDs vs. traditional retailers (%)	.47	.80	-1.11	2.04	.56
Grocery expenditure at primary chain	-.58	.48	-1.53	.37	.23
Number of shopping trips to primary chain	.00	.02	-.03	.04	.90

Notes: HD = hard discounter. Number of treated (control) households: 411 (12,937) for variables unconditional on a trip observed in a week, and 163 (5,299) for variables conditional on a trip observed in a week.

B: HOUSEHOLDS WITH MIDDLE INCOME

	Estimate		95% Confidence Interval		
	$\hat{\tau}$	Jackknife SE	Lower Bound	Upper Bound	*p*-Value
Total grocery expenditure (in euros)	.75	.78	-.77	2.27	.34
Total number of units purchased	.23	.70	-1.14	1.60	.74
Share of private labels (%)	-.95	.62	-2.17	.26	.12
Share of purchases on promotion (%)	-.54	.37	-1.28	.19	.15
Number of shopping trips	.00	.03	-.06	.06	.97
Number of different retail chains visited	-.01	.02	-.05	.03	.59
Trip share HDs vs. traditional retailers (%)	-.48	.89	-2.23	1.28	.59
Grocery expenditure at primary chain	.46	.72	-.96	1.87	.53
Number of shopping trips to primary chain	.00	.02	-.05	.04	.81

Notes: HD = hard discounter. Number of treated (control) households: 375 (15,818) for variables unconditional on a trip observed in a week, and 154 (6,705) for variables conditional on a trip observed in a week.

TABLE WA14 (CONTINUED)
C: HOUSEHOLDS WITH HIGH INCOME

	Estimate		95% Confidence Interval		
	$\hat{\tau}$	Jackknife SE	Lower Bound	Upper Bound	*p*-Value
Total grocery expenditure (in euros)	1.19	1.34	-1.45	3.82	.38
Total number of units purchased	1.02	1.06	-1.05	3.10	.33
Share of private labels (%)	-.13	1.07	-2.23	1.96	.90
Share of purchases on promotion (%)	-.58	.74	-2.04	.87	.43
Number of shopping trips	.02	.05	-.07	.11	.68
Number of different retail chains visited	-.04	.04	-.12	.03	.27
Trip share HDs vs. traditional retailers (%)	-.44	.74	-1.90	1.01	.55
Grocery expenditure at primary chain	1.66	1.30	-.88	4.20	.20
Number of shopping trips to primary chain	.07	.04	-.00	.14	.05

Notes: HD = hard discounter. Number of treated (control) households: 106 (3,548) for variables unconditional on a trip observed in a week, and 43 (1,470) for variables conditional on a trip observed in a week.

TABLE WA15: SYNTHDID ESTIMATES OF THE IMPACT OF INTRODUCING A BAN ACROSS AGE GROUPS

A: HOUSEHOLDS AGED LOWER THAN 65

	Estimate		95% Confidence Interval		
	$\hat{\tau}$	Jackknife SE	Lower Bound	Upper Bound	*p*-Value
Total grocery expenditure (in euros)	.51	.54	-.55	1.57	.35
Total number of units purchased	.09	.47	-.83	1.01	.84
Share of private labels (%)	-.28	.52	-1.30	.73	.58
Share of purchases on promotion (%)	-.10	.34	-.76	.56	.77
Number of shopping trips	.01	.02	-.03	.05	.63
Number of different retail chains visited	.00	.02	-.04	.03	.86
Trip share HDs vs. traditional retailers (%)	-.36	.75	-1.83	1.10	.63
Grocery expenditure at primary chain	.32	.51	-.68	1.31	.53
Number of shopping trips to primary chain	.00	.02	-.03	.03	.93

Notes: HD = hard discounter. Number of treated (control) households: 646 (24,286) for variables unconditional on a trip observed in a week, and 237 (9,679) for variables conditional on a trip observed in a week.

B: HOUSEHOLDS AGED 65 OR HIGHER

	Estimate		95% Confidence Interval		
	$\hat{\tau}$	Jackknife SE	Lower Bound	Upper Bound	*p*-Value
Total grocery expenditure (in euros)	.22	.63	-1.02	1.47	.72
Total number of units purchased	.03	.45	-.85	.92	.94
Share of private labels (%)	.07	.55	-1.01	1.16	.89
Share of purchases on promotion (%)	-.90	.43	-1.74	-.07	.03
Number of shopping trips	.01	.03	-.05	.07	.72
Number of different retail chains visited	-.01	.02	-.06	.03	.54
Trip share HDs vs. traditional retailers (%)	.52	.58	-.61	1.65	.36
Grocery expenditure at primary chain	.07	.62	-1.14	1.29	.90
Number of shopping trips to primary chain	.01	.02	-.04	.05	.77

Notes: HD = hard discounter. Number of treated (control) households: 246 (8,017) for variables unconditional on a trip observed in a week, and 124 (3,795) for variables conditional on a trip observed in a week.

TABLE WA16: SYNTHDID ESTIMATES OF THE IMPACT OF INTRODUCING A BAN ACROSS HOUSEHOLDS THAT VISIT A BELOW VS. ABOVE MEDIAN AMOUNT OF RETAIL CHAINS

A: HOUSEHOLDS THAT VISIT A BELOW MEDIAN AMOUNT OF RETAIL CHAINS

	Estimate		95% Confidence Interval		
	$\hat{\tau}$	Jackknife SE	Lower Bound	Upper Bound	*p*-Value
Total grocery expenditure (in euros)	.40	.54	-.66	1.45	.46
Total number of units purchased	.11	.42	-.72	.94	.80
Share of private labels (%)	.12	.59	-1.03	1.28	.84
Share of purchases on promotion (%)	-.03	.37	-.76	.69	.93
Number of shopping trips	.01	.02	-.03	.04	.69
Number of different retail chains visited	-.02	.01	-.05	.01	.13
Trip share HDs vs. traditional retailers (%)	-.47	.82	-2.07	1.12	.56
Grocery expenditure at primary chain	.31	.54	-.75	1.37	.57
Number of shopping trips to primary chain	.00	.02	-.03	.04	.79

Notes: HD = hard discounter. Number of treated (control) households: 510 (16,130) for variables unconditional on a trip observed in a week, and 163 (4,940) for variables conditional on a trip observed in a week.

B: HOUSEHOLDS THAT VISIT AN ABOVE MEDIAN AMOUNT OF RETAIL CHAINS

	Estimate		95% Confidence Interval		
	$\hat{\tau}$	Jackknife SE	Lower Bound	Upper Bound	*p*-Value
Total grocery expenditure (in euros)	.31	.72	-1.10	1.72	.66
Total number of units purchased	-.01	.64	-1.25	1.24	.99
Share of private labels (%)	-.66	.52	-1.67	.36	.21
Share of purchases on promotion (%)	-.69	.37	-1.41	.03	.06
Number of shopping trips	.00	.04	-.06	.07	.89
Number of different retail chains visited	.00	.02	-.04	.05	.89
Trip share HDs vs. traditional retailers (%)	.13	.67	-1.19	1.45	.84
Grocery expenditure at primary chain	-.07	.61	-1.26	1.12	.91
Number of shopping trips to primary chain	.01	.02	-.04	.05	.80

Notes: HD = hard discounter. Number of treated (control) households: 382 (16,173) for variables unconditional on a trip observed in a week, and 198 (8,534) for variables conditional on a trip observed in a week.

TABLE WA17: SYNTHDID ESTIMATES OF THE IMPACT OF INTRODUCING A BAN ACROSS LOW, MODERATE, AND HIGH PRICE SENSITIVE HOUSEHOLDS

A: LOW PRICE SENSITIVE HOUSEHOLDS

	Estimate		95% Confidence Interval		
	$\hat{\tau}$	Jackknife SE	Lower Bound	Upper Bound	*p*-Value
Total grocery expenditure (in euros)	-.28	1.13	-2.49	1.93	.81
Total number of units purchased	.27	.92	-1.55	2.08	.77
Share of private labels (%)	.32	1.12	-1.87	2.50	.78
Share of purchases on promotion (%)	-1.05	.73	-2.49	.38	.15
Number of shopping trips	-.03	.04	-.12	.05	.41
Number of different retail chains visited	-.05	.03	-.12	.01	.12
Trip share HDs vs. traditional retailers (%)	-.19	.92	-2.00	1.63	.84
Grocery expenditure at primary chain	.08	1.19	-2.26	2.41	.95
Number of shopping trips to primary chain	-.02	.04	-.09	.06	.66

Notes: HD = hard discounter. Number of treated (control) households: 116 (3,892) for variables unconditional on a trip observed in a week, and 52 (1,582) for variables conditional on a trip observed in a week.

B: MODERATE PRICE SENSITIVE HOUSEHOLDS

	Estimate		95% Confidence Interval		
	$\hat{\tau}$	Jackknife SE	Lower Bound	Upper Bound	*p*-Value
Total grocery expenditure (in euros)	.65	.83	-.97	2.27	.43
Total number of units purchased	-.37	.65	-1.65	.90	.57
Share of private labels (%)	-.73	.80	-2.30	.84	.36
Share of purchases on promotion (%)	-1.28	.56	-2.38	-.18	.02
Number of shopping trips	.06	.03	.00	.13	.06
Number of different retail chains visited	.01	.03	-.04	.06	.64
Trip share HDs vs. traditional retailers (%)	-.25	.64	-1.50	1.01	.70
Grocery expenditure at primary chain	-.54	.78	-2.06	.98	.49
Number of shopping trips to primary chain	.02	.02	-.03	.06	.51

Notes: HD = hard discounter. Number of treated (control) households: 215 (7,931) for variables unconditional on a trip observed in a week, and 87 (3,210) for variables conditional on a trip observed in a week.

TABLE WA17 (CONTINUED)
C: HIGH PRICE SENSITIVE HOUSEHOLDS

	Estimate		95% Confidence Interval		
	$\hat{\tau}$	Jackknife SE	Lower Bound	Upper Bound	*p*-Value
Total grocery expenditure (in euros)	.32	.53	-.72	1.35	.55
Total number of units purchased	.28	.43	-.56	1.12	.52
Share of private labels (%)	-.20	.51	-1.20	.81	.70
Share of purchases on promotion (%)	.04	.34	-.63	.71	.90
Number of shopping trips	.00	.03	-.05	.05	.86
Number of different retail chains visited	.01	.02	-.03	.04	.75
Trip share HDs vs. traditional retailers (%)	-.04	.84	-1.69	1.61	.96
Grocery expenditure at primary chain	.54	.50	-.44	1.52	.28
Number of shopping trips to primary chain	.00	.02	-.04	.03	.96

Notes: HD = hard discounter. Number of treated (control) households: 521 (19,214) for variables unconditional on a trip observed in a week, and 206 (8,256) for variables conditional on a trip observed in a week.

WEB APPENDIX K: READERSHIP OF DIGITAL STORE FLYERS

To understand readership of digital store circulars, we surveyed households in treatment and control municipalities in January 2021, just after the introduction of the ban in the last six municipalities, to ask them whether they use digital store flyers (yes/no) and, if so, when they started doing so (1 month ago, 2–3 months ago, etc.), and how often (more than once a week, once a week, etc.). Table WA18, Panel A, contains the descriptive statistics for households in the control and treatment municipalities; Table WA18, Panel B, specifies these descriptive statistics for Rotterdam, the municipality that implemented a ban in November 2020, two months before our survey.

TABLE WA18: DESCRIPTIVE STATISTICS READERSHIP OF DIGITAL STORE FLYERS

A: HOUSEHOLDS IN TREATMENT MUNICIPALITIES VS. CONTROL MUNICIPALITIES

		Treated (N = 777)	Control (N = 1,409)	Chi sq.	*p*-Value
Readership:	Yes	68.7%	67.1%	.58	.45
	No	31.3%	32.9%		
If yes, since when:	1 month	3.4%	4.1%	7.89	.10
	2-3 months	9.9%	6.5%		
	4-6 months	9.9%	11.6%		
	7-12 months	11.4%	13.4%		
	Longer than 1 year ago	65.4%	64.3%		
If yes, how often:	More than once a week	18.2%	15.2%	7.17	.13
	Once a week	42.1%	40.5%		
	Multiple times a month	22.7%	21.7%		
	Once a month	9.7%	12.8%		
	Less than once a month	7.3%	9.7%		

B: HOUSEHOLDS IN ROTTERDAM VS. CONTROL MUNICIPALITIES

		Rotterdam (N = 191)	Control (N = 1,409)	Chi sq.	*p*-Value
Readership:	Yes	71.2%	67.1%	1.27	.26
	No	28.8%	32.9%		
If yes, since when:	1 month	5.1%	4.1%	4.65	.32
	2-3 months	9.6%	6.5%		
	4-6 months	11.8%	11.6%		
	7-12 months	8.1%	13.4%		
	Longer than 1 year ago	65.4%	64.3%		
If yes, how often:	More than once a week	20.6%	15.2%	6.88	.14
	Once a week	45.6%	40.5%		
	Multiple times a month	19.1%	21.7%		
	Once a month	7.4%	12.8%		
	Less than once a month	7.4%	9.7%		

Next, Table WA19 specifies the descriptive statistics per subgroup of treated households: those that do not opt in (compliers), those that opt in (defiers), and those that already opted out prior to the ban (always-takers).

TABLE WA19: DESCRIPTIVE STATISTICS READERSHIP OF DIGITAL STORE FLYERS PER SUBGROUP OF TREATED HOUSEHOLDS

		Compliers (without a sticker) (N = 286)	**Defiers (opted in) (N = 277)**	**Chi sq.[a]**	***p*-Value[a]**	**Always-takers (already opted out before) (N = 214)**	**Chi sq.[b]**	***p*-Value[b]**
Readership:	Yes	72.7%	70.0%	.50	.48	61.7%	7.29	.03
	No	27.3%	30.0%			38.3%		
If yes, since when:	1 month	3.8%	3.6%	1.35	.85	2.3%	16.14	.04
	2-3 months	12.0%	10.3%			6.1%		
	4-6 months	10.6%	13.4%			3.8%		
	7-12 months	11.1%	12.9%			9.8%		
	Longer than 1 year ago	62.5%	59.8%			78.0%		
If yes, how often:	More than once a week	18.8%	20.6%	1.00	.91	13.6%	10.24	.25
	Once a week	40.9%	40.7%			46.2%		
	Multiple times a month	24.0%	23.7%			18.9%		
	Once a month	9.6%	10.3%			9.1%		
	Less than once a month	6.7%	4.6%			12.1%		

[a] Chi-square test of the cross-tabulation between the variable in the first column against the first two subgroups of treated households.
[b] Similar as above but against all three subgroups of treated households.

REFERENCES NOT IN MAIN TEXT

Burlig, Fiona, Louis Preonas, and Matt Woerman (2020), "Panel Data and Experimental Design," *Journal of Development Economics*, 144, 102458.

Schochet, Peter Z. (2022), "Statistical Power for Estimating Treatment Effects Using Difference-in-Differences and Comparative Interrupted Time Series Estimators With Variation in Treatment Timing," *Journal of Educational and Behavioral Statistics*, 47 (4), 367-405.

Stekhoven, Daniel. J., and Peter Buehlmann (2012), "MissForest—Nonparametric Missing Value Imputation for Mixed-Type Data," *Bioinformatics*, 28 (1), 112-118.