

FOR 03 2018

ISSN: 1500-4066

March 2018

Discussion paper

What Do You Buy When No One's Watching? The Effect of Self-Service Checkouts on the Composition of Sales in Retail

BY
Andreas Olden

What Do You Buy When No One's Watching? The Effect of Self-Service Checkouts on the Composition of Sales in Retail. *

Andreas Olden †

This draft: February, 2018

Abstract

Buying items that are unhealthy or are of a private nature may carry a stigma and cause embarrassment. I analyze whether the anonymity provided by self-service checkouts changes customers' shopping patterns in grocery stores. I look at a natural experiment where two stores in a grocery-chain implement self-service checkouts. Using a triple difference estimator, comparing the sales of stigma items to the sales of mundane items and to the sales of a group of control stores, I find that the sales of stigma items increase by 10-15 percent. The increase comes from the product categories candy, chips, soda, ready-made food and alcohol. I find that the increase is caused by existing customers buying more, rather than from self-service checkouts changing the customer base. However, fully converting to self-service seems to scare away some customers and decreases overall sales.

Keywords: social friction, stigma, grocery markets, automatization, triple difference

JEL Codes: D12, D90, M20, L81

*This project has many to thank for its current form. I would especially like to thank Jarle Møen, Mathias Ekstrøm, Einar Blix Huseby and Alexander Cappelen for invaluable feedback. The project has been presented at 68 Degrees North Conference on Behavioral and Experimental Economics (2017) and the Annual Congress of the European Economic Association (EEA) in Lisbon (2017). I thank the conference attendees for their feedback.

†The Department of Business and Management Science, NHH Norwegian School of Economics
Norwegian Centre for Taxation, NHH
FAIR The Choice Lab, NHH
Email: andreas.olden@nhh.no

1 Introduction

Imagine standing in line at your local convenience store. Your basket contains frozen pizza, soda and chocolate for Friday movie night. Maybe you are planning to purchase candy, condoms or tampons. The content of this basket is not something you want to share with others. Yet, in many purchasing situations, this is what you are forced to do. First, you stand in line with people around, then you have to display your items at the registry for scanning. Having your items displayed like this may cause a mild embarrassment. Goldfarb et al. (2015) call this embarrassment a social friction, and show that it can change your behavior. Dahl et al. (2001) show that even an imagined social presence may cause embarrassment, while Ariely and Levav (2000) argue that in social settings, people choose options that undermine their personal preferences because of self-presentation goals.

Self-service checkouts allow you to move items anonymously from basket to bag, removing the social component. This should increase the sales of items prone to social frictions. I suggest two reasons for why a product may be a stigma product and cause a social friction; first that the item is unhealthy, second that the consumption of the item is of a private nature. To investigate these mechanisms, I look at a natural experiment where two stores in a grocery-chain implements self-service checkouts.

The simplest approach to check if self-service increased sales of stigma items would be to compare the sales before and after the systems were introduced. However, this may pick up general sales trends. I control for this by comparing sales growth to the sales growth of control stores that did not introduce self-service. Another problem is that self-service checkouts may change sales in other ways than through reduced friction, for instance, by making the store more efficient, thus increasing all sales. To control for such effects I compare the changes in the stigma category to changes in a reference category of items that are not stigmatized. To incorporate both controls in one estimation framework, I use the triple difference estimator.

All stores are in urban Scandinavian areas, and in the same country. I have daily information on quantities sold by detailed product codes. The chain coordinates prices and campaigns nationally for all the products used in the analysis, and all products used are available in all stores. The data spans more than four years, starting in January 2010. Treatment store 1 introduced the system in the summer of 2011 and kept no traditional cash registers. Treatment store 2 introduced the system the summer of 2012, but kept one manned cash register. Both of the stores that introduced self-service introduced the systems at, or close to the original registers, and no major rebuilding occurred.

The estimate for the overall effect of the sales of stigma items is a 10-15 percent increase. This corresponds to an additional sale of 150-230 units per day in the stores analyzed, and

is driven mostly by candy, chips, soda, ready-made food, and alcohol. Considering the size of the estimates, knowledge of these types of effects may increase profit directly through the investment decision, or imply re-optimizing the product-cost cross-subsidization. Also, due to the size of the market and the nature of the products, the findings may warrant public policy responses. The analysis is restricted to the grocery market, but the suggested mechanisms imply that there are other markets that may experience similar frictions, in particular pharmacies.

2 Literature Review

There is a relatively large literature in psychology and other social sciences that document a feeling of embarrassment that arises relating to different situations and products. However, those studies do not document changes in real behavior. One exception is Goldfarb et al. (2015) who considers two case studies where retailers have changed the retail format towards less social interaction. They find that in both cases, there is an increase in the sale of items that are possibly embarrassing to buy, and they attribute this increase to the more anonymous purchasing situation.¹

The first case study in Goldfarb et al. (2015) is a field experiment in Systembolaget, Sweden's government-run alcohol retail monopoly, conducted in the 1990s. Systembolaget has a monopoly on retailing beverages with more than 3.5 % alcohol. Initially, all items were behind a manned counter, but then some of the stores were allowed to convert to self-services stores. The stores that converted experienced an increase in sales, and the increase was larger for items with names that were hard to pronounce. The authors attribute this increase to the removal of the customers' fear of seeming unsophisticated by pronouncing the names incorrectly. The study has two main limitations. First, the authors cannot rule out that the increase is just reduced search costs, meaning that it became easier and less costly for customers to view the full selection. Second, they cannot rule out that customers avoid items with difficult names to avoid misunderstandings with the clerk. My study does not suffer from these issues, as the items do not change place and the social interaction is visual, and not through dialogue.

The second case study in Goldfarb et al. (2015) is a pizza chain that introduced online orders. The main finding is that relative to phone orders, online orders increased the number of ingredients and calories. The authors attribute this to people not wanting to seem finicky, ordering strange combinations or toppings, and that they are afraid of seeming unhealthy ordering items such as extra cheese.

¹See Goldfarb et al. (2015) for a more comprehensive literature review of social stigma.

The disadvantage of the second case study is that there may be unobserved selection of customers into the web platform, which they cannot completely rule out. In addition, ordering over a phone may induce a feeling of urgency due to the social interaction with the clerk, while the web platform allows for time to reflect. This may affect the orders. While they do control for the role of miscommunication by looking at the errors in the orders, miscommunication may also be an issue. In my study, these problems are not relevant, as there is no time pressure when shopping, and there are limited miscommunication possibilities.

By addressing several of the limitations in Goldfarb et al. (2015), I substantiate that social interaction has real consequences in retail. Studying the grocery market, I also provide evidence from a large market that uses a different physical infrastructure from that of Goldfarb et al. (2015).

Other related works include Dahl et al. (1998) who suggests that buying condoms is embarrassing, and that this embarrassment may reduce condom sales and have negative public health consequences. Furthermore, they also show that embarrassed people have a tendency to buy from vending machines. Dabholkar et al. (2003) also find some evidence that people use self-service to avoid social interaction.

3 Data and Variable Definitions

3.1 Data and Data Cleaning

The dataset contains daily sales in quantities of all products in 14 grocery stores, from January 2010 to June 2014. Treatment store 1 introduced self-service in the summer of 2011, while treatment store 2 introduced self-service in the summer of 2012. In total, there are about 27 million observations. The analyses are conducted on subsets of these data, containing about 38 000 observations, where sales are aggregated to the daily sales per store, but split on stigma items and reference items.

I want items that are available in all stores and have therefore removed locally produced items. Also, energy bars were removed from cookies, and semi-finished food products were removed from ready-made food. This is because I believe that they are not stigmatized or significantly less stigmatized than the other items in the same category. Only positive sales were included, but the negative observations were few in number, and probably stemming from inventory. I have chosen to err on the inclusive side when choosing items, meaning that the categories provided by the grocery chain remain mostly intact.

All product categories are presented in Table 1. The table distinguishes between items with and without a minimum age limit for purchase, whether they are chosen based on the

unhealthy criterion or the private-consumption criterion, and whether it is a stigma category or a reference category.

There is no price data available. Fortunately, prices and campaigns are nationally coordinated, so all stores have the same prices. Among the chosen categories, the only exceptions are items that are about to expire, but these items are likely to be few in numbers and unrelated to other store characteristics.

3.2 The Stores

There are 14 stores in the dataset, and they are all located in medium to large cities within a Scandinavian country. There are two treatment stores. Treatment store 1 did not keep any manned cash registers, but converted fully to the automated option. Treatment store 2 kept one manned cash register. The different implementation strategies may have different effects, which warrants a distinction between them in the analysis. The two stores did not change much except for this conversion, and the automated option is either at, or close to the original location of the manned cash registers. The self-service checkout area is self-contained and has several automated registers, in which you place your basket on one side of a scanner, you pick up an item, scan it, and put it straight into a plastic bag on the other side.

Both stores automatically calls on an employee if you scan a product with an age limit. The employee then comes over, checks your ID, and leaves again. This introduces some social interaction, and for this reason I make a distinction between items with and without an age limit. There may also be people that are uncertain about how the age verification process works, and avoid self-service when buying items with an age limit, which again might yield interactions with whether the store converts fully. Based on this, one might expect the stigma effect to be smaller for items with an age limit. However, many of the items that are the most likely to be stigmatized, such as cigarettes, have an age limit, so the total effect is unclear.

As of late 2013, it became possible to register your fingerprint for age verification purposes. Since I do not have data on the implementation, and it only affects a short period at the end of the dataset, I do not analyze it explicitly. However, one should expect that this technology increases the sales of stigma items with an age limit.

3.3 The Market

The stores have a recognized brand name, and they are well known in the regions they operate. Their market is urban, inhabited areas, competing mostly for every day grocery shopping. Consumer research for the country in question shows that the most important factor for choosing grocery stores is proximity to home and quality of products. These findings are

stable over time (Lavik and Schjoll, 2012). The survey by Lavik and Schjoll (2012) shows that 55 percent of the respondents ranks proximity to home as very important, and 83 percent use the closest store for their week-day shopping. Also, only 6 percent use more than 3 stores regularly, and 72 percent use only 1 or 2 stores. The market is further characterized by 57 percent of the customers shopping at least 3 times per week. Taken together, this implies a market that is strongly proximity- and habit driven, with many repeat customers going to the same stores, close to home.

3.4 Selection of the Stigma Categories

Unhealthiness as a marker of stigma was chosen for many reasons. First of all, there seems to be strong indicators that people believe that “you are what you eat” (Stein and Nemeroff, 1995), and that you are judged negatively for eating unhealthy foods. Second, it seems that stereotypes concerning food are used for impression management (Vartanian et al., 2007). Third, people know what is healthy or not according to public dietary guidelines (Povey et al., 1998).

Items of a private nature are somewhat more difficult to define. There is a broad range of items that may have some stigma attached to them. Dahl et al. (1998) suggests that buying condoms is embarrassing, while (George and Murcott, 1992) argues that products related to female hygiene are stigmatized. Pharmaceuticals are arguably also products of a private nature.

3.5 Selection of the Reference Categories

There are several reasonable expectations for sales changes when a store introduces self-service checkouts. The normal cash registers may be bottlenecks. Removing them makes the store more efficient, attracting more customers and increases all sales. It is also possible that introducing new technology may scare away some customers. In order to isolate the effect on stigma sales caused by the new and anonymous purchasing situation, I compare the sales of stigma items to a reference category consisting of milk, child related items, and laundry and cleaning items. The criteria for choosing reference items is that; they are not stigmatized, common, mundane, have a stable sales pattern, and are sold in all stores. This gives a representative picture of the customer base that visits each store, as well as an indicator of the effects on the composition of sales.

Milk is my preferred reference category. I include all normal types of milk. This ranges from whole milk to skimmed milk, milk with and without lactose, and also all ecological versions of these products. Milk is clearly not stigmatized and has a short shelf-life, which

ensures a high turnover rate and stable sales. I also believe that most customers occasionally buy milk, which ensures representativeness and high sales. The second reference category is child related items. It includes diapers, children’s food and other children’s apparel such as pacifiers. I believe these items also fulfill all my criteria. The third reference category is laundry and cleaning items. It includes laundry detergents and equipment for cleaning houses and clothing. The category resembles milk in all but shelf-life, and it is something that all customers buy at least occasionally. The category is bigger than child related items, but still much smaller than milk.

Table 1: Classification of Items

	Unhealthy	Personal	Reference
No age limit	Icecream, candy, flavored milk, cookies, cakes, ready-made food 1, ready-made food 2, soda, chips	Female hygiene and pantyliners	Milk, child related items, laundry and cleaning related items
18 year limit	Alcohol, cigarettes, other tobacco	Pharmaceuticals (incl. condoms)	-

Note: Relative to ready-made food 1, ready-made food 2 is a more narrow product range of microwaveable food for a single person. Pharmaceuticals contains products both with and without an age limit. Perhaps the most interesting category of stigma items is condoms. It is captured on the pharmaceuticals category, but has so few sales that it is not meaningful to conduct analysis on it by itself.

3.6 Potential Limitations

A potential worry with the identification strategy is that some of the measured effect may come from store rebranding, and not stigma. Simply put, the store change may attract or scare away a particular type of customer based on their perception of the store. This may in turn change the composition of items in the average basket. This will not bias the coefficient estimates, but it will change the interpretation. The sales gain or loss from converting to self-service will then only apply to the early adapters of the technology, and would be driven by substitution between competing stores. Given the market structure described in section 3.3, with proximity to home and product quality being the most important factors for the choice of grocery store, I do not believe this is an important issue, particularly not for store 2 which keeps a manned cash register. In store 1 however, some customers might react to only being able to use the automated option. I return to this issue when discussing the empirical findings.

4 Identification Strategy

4.1 Triple Differences

I use a triple difference estimator in order to account for both sales trends that affect all stores, and effects that are not due to social frictions, such as increased efficiency. The estimator is presented in Equation 1. One way to view the triple difference in this setting is that it produces one difference-in-differences for the reference items, one difference-in-difference for the stigma items, and then takes the difference between the two. The final difference estimate, capturing the additional sales coming from stigma items relative to the reference items, while accounting for counterfactual trends in the control stores, is given by the coefficients on $TreatmentStore_s * TreatmentPeriod_t * StigmaProduct_i$ in Equation 1. $TreatmentStore$ is a treatment store specific dummy variable, $TreatmentPeriod$ is a treatment period specific dummy variable, $StigmaProduct$ is a stigma product dummy variable. The s , t , and i denotes store number, time period and product category. The α 's are store fixed effects. Unless otherwise specified, the outcome variable is log-transformed daily sales. This gives the coefficients an approximate percentage increase interpretation.

$$\begin{aligned} \text{Log}(Qty_{sti}) = & \alpha_s + \text{controls} + \\ & \beta_1(TreatmentStore_s * TreatmentPeriod_t * StigmaProduct_i) + \\ & \beta_2(TreatmentStore_s * TreatmentPeriod_t) + \\ & \beta_3(TreatmentStore_s * StigmaProduct_i) + \\ & \beta_4(TreatmentPeriod_t * StigmaProduct_i) + \\ & \beta_6 TreatmentPeriod_t + \beta_7 StigmaProduct_i + \epsilon_{sti} \end{aligned} \quad (1)$$

The control variables included are year and month fixed effects, and store-stigma interacted fixed effects which accounts for heterogeneity in product category size. There are also interactions variables between the treatment store dummies and dummies for the last two months prior to implementation, and then for the three months after implementation. The final pre and post controls allow for fire sales, rebuilding, implementation issues and some response time for customer adoption.

It is well known that difference-in-difference type estimations with few clusters can have significantly biased error terms (Bertrand et al., 2004). According to Cameron and Miller (2015) this can be accounted for by using the Wild Cluster bootstrap with Rademacher

weights.² I implement this procedure across all my analyses.

5 Analysis

5.1 Descriptive Statistics and Parallel Trends

Descriptive statistics are presented in Table 2, and the development over time of the most important outcome variables are shown in Figure 1.

From Table 2 we see that there are clear level differences between the various stores and product categories. This is to be expected. Total sales of stigma products ranges from a daily average of 1207 to 1579, depending on whether we look at one of the treatment stores or at the average of the control stores. The average daily sales of the reference items range from 106 to 221.

For a difference-in-differences estimation to have a causal interpretation, we need the parallel trend assumption to hold (Angrist and Pischke, 2008). The parallel trend assumption states that in the absence of treatment, the treatment group would have had the same sales trend as the control group. The assumption is typically validated by comparing the pre-treatment trend of the treatment stores with the pre-treatment trend of the control stores. The more similar they are, the more reasonable the assumption is.

The triple difference estimator is in essence, the difference between the estimated effect from two different difference-in-difference analyses. However, rather than requiring a parallel trend for each of the two categories, it only requires that the difference between the two categories follows a parallel trend. To see why, imagine a general sales increase affecting only the treatment stores. Individually, this would bias both the difference-in-difference estimator for the stigma items, and the estimator for the reference items. However, since they are both affected in the same way, deducting the first difference-in-difference from the second, the final difference will not contain the general sales increase, and therefore not be biased.

Figure 1 shows the development over time for the most important product categories, decomposed to the mean of the control stores, and for each of the two treatment stores. The first insight is that there is substantial seasonality. We also see level differences between the stores, and a large drop around implementation for treatment store 2. The regression analysis controls for all these issues.

The vertical lines in Figure 1 show when the two treatment stores introduced self-service checkouts. Evaluating the trend before the implementation of self-service gives some cause

²The procedure is described in Cameron and Miller (2015) and available in several R packages. Simply told, it transforms clustered residuals by doing a random draw from the values $[-1, 1]$, each with probability $p=0.5$, and multiplying it with the residuals.

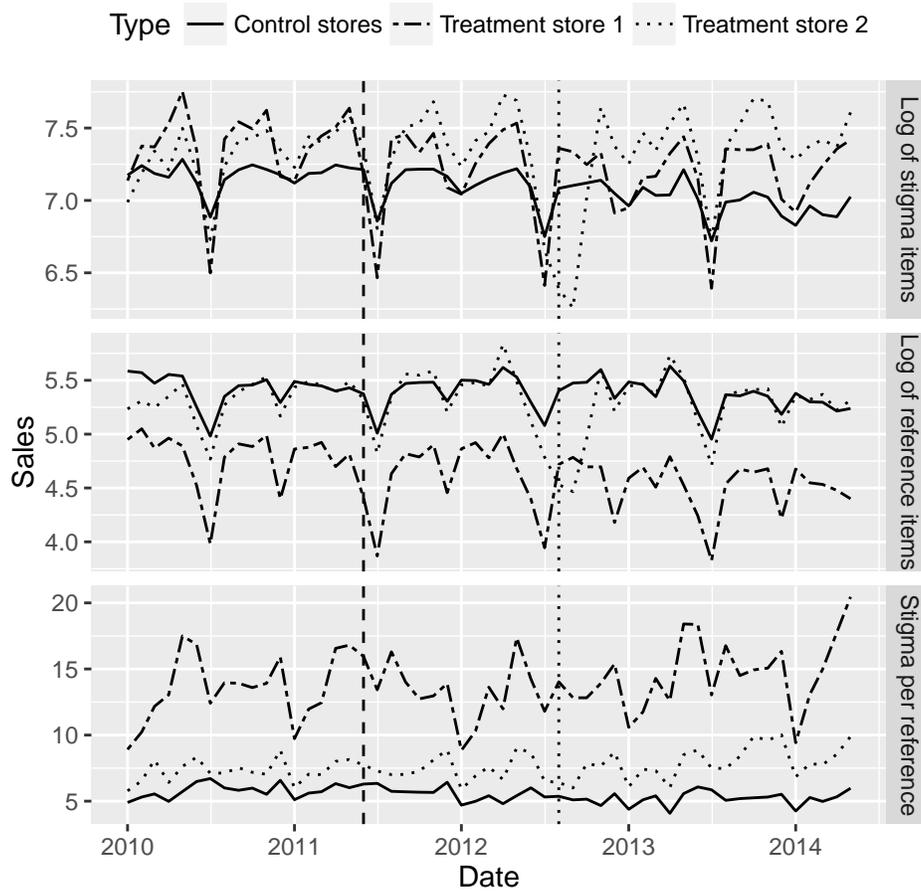
for concern. For treatment store 2, there seems to be an increase in the log-stigma sales leading up to the implementation of self-service. This increase is not seen in the control stores. However, the reference category in treatment store 2 exhibits the same behavior. As mentioned, what is important is that the difference between the stigma items and the reference items follow a parallel trend. To further underline this point we can look at the third graph from the top, which shows the development of the sales of stigma items relative to the sales of the reference items. It does not give cause for concern since it has a development over time that is very similar to the control stores. This suggests that the trends in the difference between the stigma items and the reference items are quite similar in the treatment and control stores.

Furthermore, the increase is in the peaks, not in the dips, and most of the observations are not from the peaks. In addition, the final peak before implementation is explicitly controlled for as it is quite possibly caused by a fire sale. Ignoring the final peak, the picture changes substantially, and it does not look as if there is differential pre-treatment trending.

Table 2: Descriptive Statistics of Quantities Sold per Day

Statistic	N	Mean	St. Dev.	Min	Max
<u>Control stores</u>					
Stigma items	1,387	1,207.2	345.7	324.8	2,400.1
Stigma without an age limit	1,387	740.4	182.2	174.7	2,129.9
Stigma with an age limit	1,387	466.8	212.0	68.0	1,439.2
Reference items	1,387	221.3	54.7	62.4	709.5
<u>Treatment store 1</u>					
Stigma items	1,343	1,456.8	671.0	0.0	4,132.0
Stigma without an age limit	1,343	999.4	450.6	0.0	2,071.0
Stigma with an age limit	1,343	457.5	331.5	0	2,773
Reference items	1,343	106.3	54.3	0	310
<u>Treatment store 2</u>					
Stigma items	1,377	1,578.6	639.0	15.0	8,556.9
Stigma without an age limit	1,377	1,010.7	434.6	6.0	7,876.9
Stigma with an age limit	1,377	567.8	333.4	4	4,896
Reference items	1,377	207.6	82.8	1	752

Figure 1: Graphical Representation of Sales



All figures are plots of monthly averages of daily sales. The control stores is the average of all twelve of them. The vertical lines are the implementation of the self-service checkouts for the treatment store with the same line type.

5.2 Results

The main empirical results are given in Table 3. From column 1 we see that introducing self-service checkouts increases the sales of stigma items relative to the reference items by 13.4 percent. This is given by the coefficient at the interaction term $STIG*TS*TP$. The estimate is highly significant. Since there are two treatment stores, this is an average of the effects in the two stores.

In column 2, I estimate the effect in the two treatment stores separately. The division is achieved by interacting the stigma dummy with dummies for each treatment store, and their respective treatment periods. The estimated effect for treatment store 1, the store that converted to self-service only, is 10.6 percent. The estimated effect for treatment store 2, the store that kept a manned register, is 15.8 percent. A back of the envelope calculation suggests that these effects correspond to a relative increase of 150 and 240 additional units sold of stigma items per day.

The coefficients from $TS1*TP1$ and $TS2*TP2$ give the estimated effects of introducing self-service on all items, excluding the stigma effect. The estimated effect for treatment store 1 is -19.5 percent, meaning that there has been a significant decrease in sales after introducing the self-service checkouts. Total sales will be moderated by the positive stigma effect, but is still negative. In treatment store 2, there is a positive, but not statistically significant estimated effect of 4.7 percent. In total, this store has experienced a positive and significant sales increase when the positive stigma effect is added.

Note that the decrease in sales in the store that converted fully to self-service is consistent with the embarrassment idea. Fear of potential embarrassment of not being able to use or understand how the self-service checkouts work may scare away customers, just as the anonymity of the purchasing situation may attract them. Further, this decrease is not seen in the store that kept a manned cash register, which also fits well. Instead of avoiding that store, a customer can simply choose the manned cash register instead of the self-service alternative.

Given the proximity- and habit driven market structure discussed in section 3.3, the general sales decrease in treatment store 1 is likely caused by the technology aversion described above, and not other effects like store rebranding. However, if the customers with technology aversion are not representative, the sales decrease makes the interpretation of the stigma sales estimate for treatment store 1 a bit uncertain, as it could be partly driven by a change in the customer base. Note however, that the estimated stigma sales effect in treatment store 1 is smaller than the estimated effect in treatment store 2, which did not experience a change in the sales of the reference category. Hence, the technology aversion effect seems to go in the opposite direction of the stigma effect, causing the stigma effect to be underestimated in treatment store 1.

There are also some additional points to make on how the effect of the self-service checkouts are generated. First, it is unlikely that the effect is generated by substitution between items. This is because I would expect the substitution to be either within stigma category, for instance by buying a regular soda instead of diet soda, or by substitution between stigma categories, for instance by switching from flavored milk to soda. While there may be such substitution, neither is captured in my analysis. This is because it is conducted on aggregated stigma sales in quantities, so that changing product, increasing content size of product and changing between categories will still only add the same quantity as the original product. This makes my approach a conservative estimate of the stigma effect.

Second, the fact that we do not see a significant general sales increase in treatment store 2, beyond the stigma effect, suggests that the self-service option has not attracted new customers. Otherwise, we should see an increase in the sales of the reference items. Hence, we may infer that the increase in sales of stigma products comes from the existing customers.

5.2.1 Robustness of Main Results

Column 3 estimates a difference-in-differences model, using the ratio of daily sales of stigma items to the daily sales of the reference items as the outcome variable. The difference-in-differences estimation provides a slightly different functional form, the modeling framework is somewhat simpler, and the parallel trend assumption is easier to assess as it directly corresponds to the third graph from the top in Figure 1, which is discussed earlier. The downside of the model is that it produces less information than the triple difference, and that extreme values might become more influential (ex. dividing by a number close to zero).

The treatment effect is now given by the interaction between the treatment stores and their respective treatment periods. The estimated effect for treatment store 1 is an increase in the ratio of stigma items to reference items of 1.5. The average ratio for the whole period is 13.7. A back of the envelope calculation holding the reference sale constant, suggests that the estimated effect corresponds to an increase of about 150 units or 11 percent. This is similar to the estimates from the main specification. For treatment store 2, the estimated increase is 1.2, which with an average ratio of 7.6 suggests an effect of a bit more than 230 units or 16 percent. This is also very close to the estimate from the main specification.

Finally, the main results are robust to several variations not presented here, including using standard cluster robust estimators, dropping extreme observations, severely reducing the post period, dropping pre and post controls, dropping store fixed effect and store-stigma fixed effects, and including post-treatment trends. Subdividing the effects into several post treatment periods show that the estimates are not entirely stable, but a positive and consistent average treatment effect exists across specifications.

5.2.2 Decomposing the Findings on Age

Table 3, column 4, is a triple difference comparing the sales of stigma items without an age limit to the sales of stigma items with an age limit. The reason for this split is that within the stigma categories, products without an age limit may be less stigmatized, suggesting a smaller effect. On the other hand, the social interaction is not completely removed when buying items with an age limit. This goes in the opposite direction and suggests a larger effect.

The estimate for treatment store 1 is a -4.6 percent smaller development in stigma items without an age limit than in stigma items with an age limit. The estimate is weakly significant. For treatment store 2, which did not convert fully, there is a positive and significant 6.2 percent increase in stigma items without an age limit relative to stigma items with an age limit. Taken together, no clear conclusion emerge, with respect to the role of age limits.

5.2.3 Decomposing the Findings on Stigma Categories

Table 4 presents triple difference estimates for each of the 14 categories of stigmatized items, all of them compared to the same reference category as in the main specification. The most interesting aspect to look at is which categories drive the main results. Defining a consistent category as one for which there is a positive and significant result for both treatment stores, the consistent categories are: *candy*, *chips*, *ready made food 1*, *ready made food 2*, *soda* and *alcohol*. A weaker consistency criterion, looking at estimates where both coefficient are positive, and at least one of them significant, allows us to add *cookies*, *flavored milk*, *ice cream* and *pantyliners*.

All but one category of items without an age limit fall under one of the two consistency definitions. The non-consistent category is cakes. Considering the four categories with an age limit, only alcohol is consistent. There are also differences between the stores. For treatment store 1, three out of four coefficients on the categories of stigma items with an age limit were positive and significant, namely alcohol, tobacco and pharmaceuticals (note that *tobacco* is tobacco products other than cigarettes). For treatment store 2, the one that kept a manned cash register, only alcohol was positive and significant. This is in line with previous results, which showed that treatment store 2 had a larger increase for stigma items without an age limit than for stigma items with an age limit, while treatment store 1 had no differential development between stigma items without and with an age limit. All together, ten out of fourteen categories are at least weakly consistent, and only cigarettes have a negative coefficient for both stores.

Table 3: The Effect of Self-Service on Daily Sales

	<i>Dependent variable:</i>			
	Log(Qty) (1)	Log(Qty) (2)	(Stig/Ref) (3)	Log(Qty) (4)
(STIG*TS*TP)	0.134*** (0.024)			
(STIG*TS1*TP1)		0.106*** (0.017)		
(STIG*TS2*TP2)		0.158*** (0.015)		
(u18*TS1*TP1)				-0.046* (0.022)
(u18*TS2*TP2)				0.054** (0.020)
(TS1*TP1)	-0.209*** (0.033)	-0.195*** (0.030)	1.489*** (0.131)	-0.059 (0.038)
(TS2*TP2)	0.059 (0.033)	0.047 (0.031)	1.153*** (0.147)	0.175*** (0.035)
N	38063	38063	19035	38062
R-sq	0.997	0.997	0.848	0.996
Method	Triple	Triple	DiD	Triple
Controls	Yes	Yes	Yes	Yes

Note: Columns 1-3 uses all stigma items and reference items. Column 4 uses only stigma items, but compares stigma items without an age limit to stigma items with an age limit. *Qty* is quantity measured as number of items sold. *STIG*, *TS*, *TP* and *u18* are dummy variables for *stigma products*, *treatment store*, *treatment period* and *stigma items without an age limit*, while 1 and 2 indicates the treatment store and the respective treatment period. Standard errors are Wild Cluster bootstrapped using Rademacher weight and 10 000 replications.

*p<0.05; **p<0.01; ***p<0.001

Table 4: Triple Difference Estimates of the Effect of Self-Service on Daily Sales for Individual Stigma Categories

		<i>Dependent variable:</i>													
		Cakes	Candy	Chips	Cookies	Flavored milk	Ice cream	Ready made1	Ready made2	Soda	Panty liner	Alcohol	Cigarettes	Tobacco	Pharmaceuticals
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
(STIG*TS1*TP1)		-.143*** (.032)	.072*** (.014)	.144*** (.034)	.445*** (.036)	.031 (.038)	.206*** (.032)	.309*** (.019)	.310*** (.049)	.102*** (.014)	.054 (.033)	.158*** (.038)	-.020 (.025)	.057*** (.016)	.091* (.037)
(STIG*TS2*TP2)		.315*** (.028)	.122*** (.015)	.138*** (.026)	.059 (.036)	.200*** (.034)	.024 (.026)	.220*** (.028)	.163*** (.040)	.167*** (.018)	.189*** (.036)	.196*** (.030)	-.193*** (.026)	-.016 (.022)	-.112*** (.034)
Mean(TreatStore1)	Yes	142	340	39	40	66	57	40	7	257	19	278	81	94	7
Mean(TreatStore2)	Yes	105	253	62	29	81	40	94	23	315	11	383	86	102	8
Mean(Controls)	Yes	61	200	70	19	24	31	73	17	233	10	322	93	54	8
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N		38051	38053	38048	38008	38036	37947	38047	37954	38051	37847	37625	38053	38050	37736
R-sq		0.993	0.995	0.989	0.99	0.992	0.977	0.994	0.989	0.995	0.987	0.99	0.995	0.994	0.987

Note: The dependent variable is the log of the number of items sold of each product category. "STIG", "TS" and "TP" are dummy variables for "stigma products", "treatment store" and "treatment period", while 1 and 2 indicates the treatment store and the respective treatment period. All estimates include the full set of triple difference variables, uses the full reference category, and have a full set of control variables. Standard errors are Wild Cluster bootstrapped using Rademacher weight and 10 000 replications.

*p<0.05; ** p<0.01; *** p<0.001

6 Conclusion

I find ample evidence suggesting that when a store converts to self-service, it sells more stigmatized items, meaning items that are private or unhealthy, relative to non-stigmatized items such as milk. I attribute this increase to the removal of social friction in the form of the purchasing anonymity provided by self-service checkouts. The estimated effect is a 10-15 percent increase in stigma items.

There are two business models to choose from when converting to self-service. Either the store can keep some manned cash registers, or it can convert to self-service only. My findings suggest that the former may be the preferred strategy. The store that converted fully experienced a net reduction in total sales compared to the control stores. There is no reduction in sales for the store that kept a manned cash register, suggesting that converting fully scares away some customers.

The estimated increase in stigma sales is most likely generated by existing customers buying more. The main reason for embarrassment seems to be that an item is unhealthy, but there is some evidence that also items of a private consumption nature causes mild embarrassment. The main drivers of the effect are the categories candy, chips, soda, ready-made food, and alcohol. Combining the size of these categories with the size of the estimates, there is good reason to take such effects into account when considering implementing self-service systems.

References

- Angrist, J. D. and J.-S. Pischke (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.
- Ariely, D. and J. Levav (2000). Sequential choice in group settings: Taking the road less traveled and less enjoyed. *Journal of Consumer Research* 27(3), 279–290.
- Bertrand, M., E. Duflo, and S. Mullainathan (2004). How much should we trust differences-in-differences estimates? *The Quarterly journal of economics* 119(1), 249–275.
- Cameron, A. C. and D. L. Miller (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources* 50(2), 317–372.
- Dabholkar, P. A., L. M. Bobbitt, and E.-J. Lee (2003). Understanding consumer motivation and behavior related to self-scanning in retailing: implications for strategy and research on technology-based self-service. *International Journal of Service Industry Management* 14(1), 59–95.
- Dahl, D. W., G. J. Gorn, and C. B. Weinberg (1998). The impact of embarrassment on condom purchase behaviour. *Canadian Journal of Public Health/Revue Canadienne de Sante'e Publique* 89(6), 368–370.
- Dahl, D. W., R. V. Manchanda, and J. J. Argo (2001). Embarrassment in consumer purchase: The roles of social presence and purchase familiarity. *Journal of Consumer Research* 28(3), 473–481.
- George, A. and A. Murcott (1992). Research note: Monthly strategies for discretion: Shopping for sanitary towels and tampons. *The Sociological Review* 40(1), 146–162.
- Goldfarb, A., R. C. McDevitt, S. Samila, and B. S. Silverman (2015). The effect of social interaction on economic transactions: Evidence from changes in two retail formats. *Management Science* 61(12), 2963–2981.
- Lavik, R. and A. Schjoll (2012). Endring i butikkstruktur og handlemønster i norsk dagligvarehandel. retrieved february 12, 2018. *SIFO* <http://www.hioa.no/om-hioa/senter-for-velferds-og-arbeidslivsforskning/sifo/publikasjoner-fra-sifo/endring-i-butikkstruktur-og-handlemoenster-i-norsk-dagligvarehandel>.
- Povey, R., M. Conner, P. Sparks, R. James, and R. Shepherd (1998). Interpretations of healthy and unhealthy eating, and implications for dietary change. *Health Education Research* 13(2), 171–183.

Stein, R. I. and C. J. Nemeroff (1995). Moral overtones of food: Judgments of others based on what they eat. *Personality and Social Psychology Bulletin* 21(5), 480–490.

Vartanian, L. R., C. P. Herman, and J. Polivy (2007). Consumption stereotypes and impression management: How you are what you eat. *Appetite* 48(3), 265–277.