The Perils of Private Provision of Public Goods

Umit G. Gurun
University of Texas at Dallas

David H. Solomon
Boston College

First version: February 10, 2019
This version: November 1, 2019

Abstract: In May 2018, in response to protests, Starbucks changed its policies nationwide to allow anybody to sit in their stores and use the bathroom without making a purchase. Using a large panel of anonymized cellphone location data, we estimate that the policy led to a 6.8% decline in store attendance at Starbucks locations relative to other nearby coffee shops and restaurants. This decline cannot be calculated from Starbucks' public disclosures, which lack the comparison group of other coffee shops. The decline in visits is around 84% larger for stores located near homeless shelters. The policy also affected the intensive margin of demand: remaining customers spent 4.2% less time in Starbucks relative to nearby coffee shops after the policy enactment. Wealthier customers reduced their visits more, but black and white customers were equally deterred. The policy led to fewer citations for public urination near Starbucks locations, but had no effect on other similar public order crimes. These results show the difficulties of companies attempting to provide public goods, as potential customers are crowded out by non-paying members of the public.

The central problem in all of economics is scarcity. Wants are essentially unlimited, but the means available to satisfy them are not. In general, markets solve this problem via the price mechanism, whereby scarce goods are allocated based on people's willingness and ability to pay. From a philosophical perspective, this has its benefits and shortcomings. Prices serve to allocate goods to people who likely value them more, and provide incentives to producers to produce more. The tradeoff is one of fairness, whereby the rich receive more. But allocating goods differently does not thereby solve the underlying shortage problem. When wants exceed goods, *someone* has to be denied – the question is merely whom. Alternative allocation mechanisms end up favoring other groups, such as those with a lower cost of their time, those with personal or political connections, or simply those who are lucky.

We study this problem in the context of how companies allocate semi-public goods like seating and bathrooms. On April 15, 2018, a Starbucks store in Philadelphia called the police after two African-American men refused to leave the store, despite not purchasing anything.[1] This led to a series of nationwide protests accusing Starbuck's existing policies of exhibiting racial bias.[2] In response, Starbucks held a day of sensitivity training for all employees on May 29, 2018[3], and announced a new nationwide policy that anyone was welcome to sit in Starbucks stores and use the bathrooms, without any need for a purchase[4].

Bathrooms are an example of a public good, but one which receives considerably less attention from economists than textbook examples like courts, roads, lighthouses, and so on. In the absence of perfect law enforcement, people urinating and defecating on the streets has clear negative externalities for those in the surrounding areas. However, those individuals or stores who do not contribute towards common bathrooms still benefit from their availability, creating a free-rider problem. Similarly, the enforcement of

---

[1] https://www.cnn.com/2018/04/14/us/philadelphia-police-starbucks-arrests/index.html

[2] https://www.nbcnews.com/business/business-news/starbucks-brand-could-see-permanent-damage-after-racism-claims-say-n866591

[3] The half day of training and closure of stores nationwide was announced on April 17, 2018. https://stories.starbucks.com/press/2018/starbucks-to-close-stores-nationwide-for-racial-bias-education-may-29/

[4] The policy change was reported in the Wall Street Journal on May 19, 2018 https://www.wsj.com/articles/starbucks-creates-policy-on-nonpaying-guests-1526745600

vagrancy laws suggests that many municipalities find the presence of people sitting or sleeping on streets to also have negative externalities on the ambience of those areas.

What is striking about the change is that Starbucks *already had* a quasi-public bathroom policy. While the enforcement varied by store, the implicit arrangement in most places seemed to be that one could effectively use the bathroom as long as one looked like they *might* be about to make a purchase (or could officially use it, for the price of a few dollars in purchases).[5] To move from a mostly open policy to a completely open policy thus affected only a relatively small fraction of the populace: those unable to credibly signal that they might be willing and able to spend a few dollars at the store. If these changes impacted the ability of other customers to use Starbucks amenities, or if customers prefer to not be around certain types of clientele, then the change could nonetheless have significant effects.

We explore this question using anonymized cellphone location data from more than 10 million devices from January 2017 to October 2018. This allows us to estimate monthly visits to each Starbucks location, for the roughly 74% of Starbucks locations across the US where GPS data can be measured reliably. We control for city-month fixed effects to ensure we are not just measuring differences in local economic conditions. We also control for consumers' overall tastes for coffee by comparing Starbucks to other nearby coffee shops (e.g. Peet's, Coffee Bean and Tea Leaf, etc.), and other local restaurants.

Our baseline empirical specification suggests that Starbucks stores experienced a 6.8% decrease in visits after the enactment of the policy, compared with similar coffee shops and restaurants.[6] This gap is fairly consistent across various specifications that control for different average levels of store visits, time trends, and changes at the census block level over time. Importantly, this decline would not have been

---

[5]This anecdotal perception mirrors remarks made by Howard Schulz around the time of the incident. From CNN: *"Earlier this month, Starbucks Chairman Howard Schultz, speaking at the Atlantic Council in Washington, teased the policy change. He said the company used to have a "loose" policy of only allowing paying customers to use the bathroom, with the decision ultimately left to the store manager, but that stores would change that rule."* https://money.cnn.com/2018/05/20/news/companies/starbucks-bathroom-policy/index.html

[6] The effect of Philadelphia events on Starbucks sales have been also noticed by financial analysts. In one of the company earnings conference calls following the event, one analyst asked *"... with respect to the same-store sales commentary, ...sometimes you've had political activism and things like that, ... have disrupted some of your urban stores and you saw a little bit of volume slow down."*.
See https://money.cnn.com/2018/05/20/news/companies/starbucks-bathroom-policy/index.html

apparent by examining only Starbucks publicly released comparable store sales changes, as these lack the comparison group of what happened to other coffee shops. In the third quarter (covering April, May and June 2018), Starbucks reported an increase in comparable store sales of 1%, and we approximate fourth quarter comparable store sales growth at 3%.[7] Over, the same two periods, we measure the raw change in cell phone visits to Starbucks as -0.1% and 2.4% (that is, relative just to the previous period's visits at Starbucks locations). This reinforces the conclusion that our cell phone visits numbers fairly closely match Starbucks' public sales disclosures. However, they also highlight the importance of evaluating the policy based on what would have been expected to occur from trends in other nearby coffee shops. After the policy change, Starbucks saw a small time-series increase in visits, whereas absent the policy a much larger increase would have been expected. Put differently, the general boom in visits to all coffee shops at the time helped disguise the fact that the new policy appears to have significantly reduced visits to Starbucks.[8]

More strikingly, the decrease after the policy enactment is significantly larger the closer the location is to a homeless shelter. Stores less than two kilometers away experienced declines of 8.3% relative to nearby coffee shops, while stores more than 10 km away experienced declines of only 4.5%. Again, this decline in attendance is not from worsening economic conditions in these areas – rather it captures the change in Starbucks relative to nearby coffee shops experiencing the same local economic conditions. The decline in far-off locations indicates that the problem is not limited to stores near homeless populations. Even increased use by the general public of bathrooms and tables is estimated to have negative impacts, partly because people who might have previously felt compelled to make a purchase in order to sit or use the bathroom now no longer do. We also find evidence that the relative decline in Starbucks visits is greater in denser urban environments, consistent with greater foot traffic also creating more demand for bathrooms. However, controlling for homeless shelter distance considerably reduces the effect of density, whereas

---

[7] As we discuss in section 1.5, Starbucks reports quarterly comparable store sales growth for the first three quarters, but only annual comparable store sales growth for the fourth quarter / annual report. We use these four numbers to approximate fourth quarter comparable sales growth.
[8] Various media reports have cited anecdotal evidence consistent with such a decline, however. See, for instance: https://www.wsj.com/articles/starbucks-opened-its-bathrooms-to-everyone-and-some-people-are-worried-1527159601

controlling for density leaves the homeless shelter effect largely unchanged. Nonetheless, because it is difficult to precisely estimate the location of the homeless population, proximity to homeless shelters may also be proxying for other aspects of the urban environment that make the effect of the policy larger, not just the effect of the homeless themselves.

To further rule out the possibility of unobserved confounding effects in our tests, we use the synthetic control method (Abadie and Gardeazabal (2003) and Abadie, Diamond, and Hainmueller (2012)). This method forms an artificial control group of non-Starbucks coffee shops that more closely matches the performance of each Starbucks location prior to the initiation of the bathroom policy. Under this method the concentration of the overall effect near homeless shelters is even greater, consistent with the importance of proximity to homeless shelters as a mechanism for the overall decline.

Additional supportive evidence for the main prediction is apparent in the duration of customer visits. If non-paying visitors are lingering at tables, or if bathrooms are crowded and dirty, customers may also be expected to spend less time in the store. Consistent with this, visitors to Starbucks reduced the amount of time they spent in the store by 4.2%, again relative to other coffee shops and restaurants.

Starbucks also experienced a significant change in the demographics of who visited the store. Relative to other coffee shops and restaurants, Starbucks saw a larger decline in visitors from relatively wealthier home locations. The estimated income of Starbucks customers declined by 0.3%, relative to changes in other coffee shops and restaurants. This is consistent with the interpretation that wealthier clientele either have stronger preferences against other visitors attracted by the policy, or have more desire to sit at stores for longer periods. Despite the racial angle on the initial controversy, we find no difference in the racial demographics of the home locations of Starbucks visitors after the policy. In other words, the new policy appears to have deterred both black and white customers in roughly equal amounts.

Finally, we directly establish the existence of a bathroom channel by examining police citations for public urination in several cities. We find a decrease in citations near Starbucks locations relative to other areas after the enactment of the policy. By contrast, a wide range of other minor public order crimes show no significant changes, and the effects are of inconsistent signs. This result is especially difficult to explain

5

by other mechanisms, as the crime in question is unusually specific in its relation to the policy change, the changes are all within the same city, and they are measured relative to common time and area fixed effects.

These results suggest that the new policy has been costly to Starbucks, particularly in locations closer to homeless shelters. It is worth emphasizing that these estimated declines in visits are net of various positive effects, such as customers being drawn to the store because of their new policy, a perception of the company pursuing racial equality, or customers coming in to use the bathroom and deciding to purchase something anyway. Indeed, the decline in total visits likely understates the effect on the number of paying customers, as it seems probable that at least some of the new visitors are now coming in to use the bathroom without making a purchase, and yet even with these included the total number of visits is lower. It remains possible that other effects may improve sales in ways that offset the decline in visitors. Starbucks may sell more per visitor, or benefit in long term reputation with employees or customers. However, our estimates of visits map fairly well to Starbucks public sales numbers, and after the policy change Starbucks customers have lower average income and spend less time in the store.

The potential loss in sales highlights the tradeoff private companies face when deciding whether to provide universally available public goods. Necessarily, as a store goes from providing for customers only, to credible potential customers, to those with no prospect of being customers, scarce store resources get consumed with less and less private return in response. More importantly, allocating these resources to non-paying members of the public can have outsized effects on sales by deterring paying customers. The cost, in other words, is actually lost sales, not just the possible additional staffing costs to keep the bathrooms clean (which occur on top of the effects we document).

The big remaining question, to which we do not have a strong answer, is how much of the decline is due to consumption of bundled goods, versus preferences over other customers. Under the first explanation, Starbucks customers are actually buying a bundle of coffee, tables to sit and relax at, and bathrooms to use. When they enter the restaurant and find the tables and bathrooms full, they are effectively getting less of the bundled goods they desire, and so do not purchase coffee either. The preferences explanation posits that customers may also have preferences over who the other visitors are at the store. In

other words, they may have a preference against being around populations attracted by the policy, such as the homeless, and avoid the store if such people are regularly there.

Our paper is related to several strands of the literature. In the public economics literature, earlier works such as Bergstrom, Blume, and Varian (1986) start with the notion that pure public goods would be undersupplied by voluntary contributions and then show how wealth distribution affects provision of the public good. They show that changes on the extensive margin, e.g. the decision of whether or not to become a contributor of the public good – are at least as important as adjustments on the intensive margin – the decision of how much to contribute. Andreoni (1990) challenges the view that private charity is a pure public good and argues for non-altruistic motives for giving (e.g. guilt, repentance, envy, sympathy, emulation, a taste for fairness). Recent experimental work studies the crowding out effect between pure altruism and these non-altruistic motives, and shows that indeed non-altruistic motives are important factors that determine giving (see, Bolton and Katok (1998), Fisman et al. (2007), Ottoni-Wilhelm, Vesterlund and Xie (2017), Yildirim (2014), Ribar and Wilhelm (2002), among others). Our paper contributes to this literature by showing the limits of non-altruistic motives. When certain public goods are bundled with sales, it is still profitable for companies to make them available, but at a certain point firms must decide how much of the public good to distribute to non-patrons who may also actively deter others' purchases.

Our study is also related to the burgeoning literature on whether (or to what extent) corporations should engage in socially responsible activities. This literature often starts with Milton Friedman's statement made in an New York Times Magazine op-ed piece suggesting managers should solely focus on profit maximization (subject to legal constraints) and then distribute all profits to shareholders who prefer fully decentralized giving (Friedman (1970)). Subsequent literature suggests firms could engage in charitable giving as an advertising tool which can shift consumer demand (Servaes and Tamayo, 2013). Others argue shareholders might prefer to use firms as a vehicle to discharge their social responsibilities (Baron, 2007). Bagnoli and Watts (2003) link the provision of a public good by firms to sales of their private goods, and show that level of private provision of the public good varies inversely with the competitiveness of the private good market and that the types of public goods provided are biased toward those for which

consumers have high participation value. Morgan and Tumlinson (2019) offer a theoretical model in which managerial contracts reflect shareholder concerns over both public goods and profits. Investors respond to firms' provision of public goods in the form of CSR activities. Along these lines, Dimson, Karakas, and Li (2015) show that institutional investors' active engagement in monitoring of ESG issues is often perceived as good news by shareholders. Edmans (2011), Chava (2014), and Ferrell, Liang, and Renneboog (2016) provide examples of mechanisms through which CSR can enhance shareholder wealth. In addition to these works, Kruger (2015) studies how stock markets react to positive and negative events concerned with a firm's corporate social responsibility, and Hartzmark and Sussman (2019) show that mutual fund investors allocate more money to sustainable funds. Our study contributes to this line of research by showing that it is not ex-ante clear whether provision of public goods leads to increased shareholder wealth, especially when the provision of public goods by profit-maximizing companies target those groups who have the lowest willingness to pay, many of whom are amongst the most vulnerable groups in society.

1. Data and sample construction

The analysis relies on three main sources of data: (1) establishment-level foot traffic, (2) homeless shelter locations, and (3) incident-level crime reports. In this section we describe these sources and outline our sample construction.

1.1. Establishment-level foot traffic

The establishment-level foot traffic is provided by SafeGraph, a company that aggregates anonymized smartphone-location data from numerous applications (e.g. local news, weather or other information) in both Apple and Android platforms to provide insights about physical places. The underlying data cover about 10% of smartphones in the United States. The raw data consists of "pings," each of which identifies the latitude and longitude of a smartphone at a moment in time. The location information can be used to understand a device's location in detail, accurate to within a few meters. SafeGraph uses an algorithm that considers a number of features (including the proximity of the pings to the establishment's

footprint, number of pings, duration between pings) to determine whether a device visited an establishment. SafeGraph then aggregates the visits to public places like Starbucks over the course of the month and provides these anonymized aggregated numbers. To further enhance privacy, SafeGraph excludes census block group information if fewer than five devices visited an establishment in a month from a given census block group. Our sample consists of establishment-level estimates of foot-traffic using reported GPS locations from participating apps, aggregated to the monthly level and spanning the 22-month period from January 2017 to October 2018. Finally, our data allows us to estimate the demographics of visitors to an establishment in a given month, such as race, income, etc. Specifically, we infer these traits for each visitor by matching the census block group the device resides in to the average income and racial composition for the block group from the 2017 American Community Survey from the Census Bureau. For each store-month's visits, we then compute the weighted-average income and race of visitors where the fractional weights are set to the percent of visits from each residential block group.

In our analysis, we consider three mutually exclusive types of establishments: *Starbucks*, *Coffee Shops*, and *Restaurants*. We focus on non-Starbucks coffee shops because these establishments constitute a reasonable control group which would not be affected by the enactment of Starbucks' policy in the same manner as the effect on Starbucks. To construct the set of non-Starbucks coffee shops, we identify establishments that sell coffee based on two criteria. First, we identify all firms in our sample with a 6-digit NAICS code of 722515 (*Snack and nonalcoholic beverage bars*). From this set, we hand-classify each firm with at least five store locations based on company name (and web search if necessary) to determine its eligibility for the *Coffee Shop* group. As our second criteria, we consider an establishment to belong to *Coffee Shop* if the a) firm's name contains the word "coffee", and b) firm's 5-digit NAICS code is 72251 (*Restaurants and other eating places*). We consider an establishment meeting either of the previous two criteria as belonging to the *Coffee Shop* group. Finally, from the set of all remaining firms with a 5-digit NAICS code of 72251, we construct the *Restaurants* group by selecting a 25% random sample. Other coffee shops are the closest counterfactual to Starbucks, since they sell a very similar product. Investigating the

foot traffic in the *Restaurants* group is interesting because it provides a check for whether there might be unusual changes in other coffee shops, rather than Starbucks itself.

There are two sample issues that are common in the type of anonymized location data we use. First, representativeness. Our first maintained assumption is that foot traffic captured by the GPS location data does not selectively exclude customers that share a certain attribute that could be correlated with the treatment effect. According to a recent Pew research, 92% (67%) of American adults own a cellphone (smartphone).[9] While 90% of cellphone owners say they "frequently" carry their phone with them, 6% say they "occasionally" have their phones with them. Only 4% say they only "rarely" or "never" have their cellphones with them. These statistics suggest GPS location data is reasonably comprehensive enough to provide a metric that can help us measure the foot traffic.

Second, the number of devices considered in the sample increases over time, primarily due to an increase in the number of smart phone applications utilizing location information. Thus, we observe an upward trend in foot traffic in all types of establishments. While we cannot identify all factors contributing to the upward trend, we can construct an inflation factor used to de-trend our foot traffic data. More precisely, we first define $totalCount_{ct}$ as the total number of visits made by all devices in core-based statistical area (CBSA) c in month t.[10] We then scale the number of visits to establishment i in CBSA c in month t by $totalCount_{cT}/totalCount_{ct}$, where T is the final month in our sample. Thus, each visit count is adjusted to an October 2018 level based on the overall growth in that city.

We note that similar anonymized cell phone location data has been used to understand movements of individuals in other contexts, such as travel to and from Thanksgiving (Chen and Rohla 2018), hurricane evacuation (Chen, Long and Rohla 2019) and neighborhood segregation (Athey, Ferguson, Gentzkow and Schmidt 2019), suggesting it is a good proxy for actual individual movements.

---

[9] https://www.pewinternet.org/2015/08/26/chapter-1-always-on-connectivity/

[10] A CBSA is a U.S. geographic area that combines of one or more counties (or equivalents) anchored by an urban center of at least 10,000 people plus adjacent counties that are socioeconomically tied to the urban center by commuting, and is determined by the Office of Management and Budget.

1. 2.   Homeless shelter locations

Our second dataset contains homeless shelter locations and addresses, from https://www.homelessshelterdirectory.org/. This dataset is meant to be representative of general homeless population locations, rather than being an exhaustive list of all shelters or places where the homeless live. We geocode the address of each shelter using a combination of the *US Census Geocoder* and *Google Maps*, yielding a set of latitude/longitude pairs. For each establishment, we then compute the distance to each shelter using the World Geodetic System 1984 (WGS84) projection with the longitudinal zone determined by the establishment's longitude, and take the minimum distance across all shelter locations.

1.3.   Incident-level crime reports

We collect the incident-level micro-data from 2016 through 2018 reported by three cities: Austin, Denver and Pittsburgh. These cities were chosen out of a larger search of all major cities that publicly report incident-level data. The important criteria for the above cities are based on them having both geocoded incidents and a fine enough category of crime reporting to allow for reasonable numbers of crimes plausibly related to public urination. The benefit of zooming into incident-level data, rather than relying on county level aggregated crime reports often used in the crime literature (e.g. Uniform Crime Reports (UCB), or National Incident Based Reporting System (NIBRS)) is that we can identify the precise location of the incidents possibly affected by the treatment we are interested in. These three cities provide the detailed information necessary to test the unintended consequences of the Starbucks announcement. However, the incidents are not described in a uniform fashion. For this reason, we hand-classify public urination related crimes only for instances where the crime description specifically references this, such as Austin's 'URINATING IN PUBLIC PLACE' descriptor. We then aggregate up instances of public urination at a monthly interval and at a geographic level of a census block group, using the centroid of each census block group to compute the distance to the nearest Starbucks. We restrict our analysis to block groups that

experience at least one instance of public urination over the period considered, yielding a final sample of 350 block groups.

1.4.   Summary Statistics

We report summary statistics for our data in Table 1. Panel A displays statistics for the key variables used in the foot traffic analysis, measured at the store-month level. The average store in our sample experiences approximately 344 visits per month from sample devices, lasting an average of 35 minutes each. Note our data does not account for all visitors to a store, simply those using a smart phone application from which our data provider obtains location data. For stores in cities that have a homeless shelter recorded in our dataset, the average store is located 7.67 km from a shelter. Finally, for the subset of stores with a sufficient number of visitors to estimate income statistics, the average household income of visitors is $71k. Panel B partitions the sample based on the three categories we study. While *Coffee Shops* and *Restaurants* are similar in their number of visitors and estimated income, *Starbucks* establishments tend to attract more visitors with a higher estimated income. Visitors to Starbucks also tend to spend less time in the store.

In terms of numbers of stores, we have non-missing visit data for 10,752 Starbucks locations, 24,800 coffee shops, and 137,745 restaurants. This is less than the 14,620 Starbucks locations that Safegraph has business listing and footprint data for.[11] The reason is that Safegraph is unable to track visits to store locations (for all store types) located inside large structures such as indoor malls, airports, and stadiums, due to GPS scattering. Strip malls are identified correctly, as of course are stand-alone locations.

Finally, Panel C reports summary statistics related to the incident-level crime data for the 350 census block groups over the 3-year period spanning 2016 to 2018. The average distance of a block group to the nearest Starbucks is 1.5 km. Moreover, urination-related incidents appear to be a relatively rare event with 0.1 events per block group-month, or slightly less than 12 per city-month in the sample.

---

[11] Starbucks 10-K from November 2018 lists 14,606 US stores, comprising 6,031 licensed stores and 8,575 company-operated stores, as of September 30, 2018 (the slightly lower number than Safegraph counts is due to Safegraph data extending beyond this date). This suggests that the Safegraph location data represents nearly all Starbucks locations, while the visit data corresponds to roughly 74% of Starbucks stores in the U.S.

Insert Table 1 here

1.5 Comparing Cell-Phone Starbucks Visits with Starbucks' Public Disclosures

Because our anonymized cell phone location data accurately measures the location of individuals, and is available at the individual establishment level, it represents a metric of corporate performance that is impossible to obtain directly from Starbucks public disclosures (let alone for private coffee shops and restaurants, which lack any public disclosures, and form an essential part of our control group).

Nonetheless, as a verification check, we compare how our aggregated numbers match up with Starbucks public disclosures in their 10-K and 10-Q annual and quarterly earnings reports, obtained from the SEC's EDGAR database. Starbucks reports quarterly revenue numbers but these will also include factors such as changes in the number of stores, expansions into geographically different areas, etc. More usefully, Starbucks also reports its own measure of percentage change in comparable store sales, rounded to a whole number of percent. It is not clear from the discussion in their disclosures whether this is quarter-on-quarter growth, or growth of this quarter relative to same-quarter sales in the previous year (we assume the former, as computing analogs of the latter would lose us one of our two years of cellphone data). Additional details of the mechanics of the same store sales growth calculation are not provided. The closest number Starbucks reports that we can directly approximate is Change in the Number of Transactions, but this unfortunately is only reported for the Americas for two quarters over our sample.

Both metrics (revenue and comparable store sales growth) are disclosed for fiscal quarters 1 to 3. For the fourth fiscal quarter, only the whole year numbers are reported, and not fiscal quarter 4 specifically. This makes interpreting fourth quarter revenues fairly straightforward, as the difference between the total and the previous three quarters. Fourth quarter same store growth numbers can only be approximated, however. We approximate them based off annual and quarterly growth numbers, plus quarterly sales revenues.

Specifically, assume that $S_{i,j}$ represents the sales of quarter i in year j. Then we have the following for relating quarterly to annual sales growth, G, assuming that growth is quarter-on-quarter or year-on-year:

$$G_{Ann,j} = \frac{\left(S_{1,j} + S_{2,j} + S_{3,j} + S_{4,j}\right) - \left(S_{1,j-1} + S_{2,j-1} + S_{3,j-1} + S_{4,j-1}\right)}{\left(S_{1,j-1} + S_{2,j-1} + S_{3,j-1} + S_{4,j-1}\right)}$$

$$= \frac{\left(S_{1,j} - S_{1,j-1}\right)}{\left(S_{1,j-1} + S_{2,j-1} + S_{3,j-1} + S_{4,j-1}\right)} + \frac{\left(S_{2,j} - S_{2,j-1}\right)}{\left(S_{1,j-1} + S_{2,j-1} + S_{3,j-1} + S_{4,j-1}\right)}$$

$$+ \frac{\left(S_{3,j} - S_{3,j-1}\right)}{\left(S_{1,j-1} + S_{2,j-1} + S_{3,j-1} + S_{4,j-1}\right)} + \frac{\left(S_{4,j} - S_{4,j-1}\right)}{\left(S_{1,j-1} + S_{2,j-1} + S_{3,j-1} + S_{4,j-1}\right)}$$

$$= G_{1,j} * \frac{S_{1,j-1}}{\left(S_{1,j-1} + S_{2,j-1} + S_{3,j-1} + S_{4,j-1}\right)} + G_{2,j} * \frac{S_{2,j-1}}{\left(S_{1,j-1} + S_{2,j-1} + S_{3,j-1} + S_{4,j-1}\right)}$$

$$+ G_{3,j} * \frac{S_{3,j-1}}{\left(S_{1,j-1} + S_{2,j-1} + S_{3,j-1} + S_{4,j-1}\right)} + G_{4,j} * \frac{S_{4,j-1}}{\left(S_{1,j-1} + S_{2,j-1} + S_{3,j-1} + S_{4,j-1}\right)}$$

Thus we have

$$G_{4,j} = \frac{\left(S_{1,j-1} + S_{2,j-1} + S_{3,j-1} + S_{4,j-1}\right)}{S_{4,j-1}}$$

$$* \left(G_{Ann,j} - G_{1,j} * \frac{S_{1,j-1}}{\left(S_{1,j-1} + S_{2,j-1} + S_{3,j-1} + S_{4,j-1}\right)} - G_{2,j} \right.$$

$$\left. * \frac{S_{2,j-1}}{\left(S_{1,j-1} + S_{2,j-1} + S_{3,j-1} + S_{4,j-1}\right)} - G_{3,j} * \frac{S_{3,j-1}}{\left(S_{1,j-1} + S_{2,j-1} + S_{3,j-1} + S_{4,j-1}\right)} \right)$$

We use aggregate revenue numbers and average same-store-growth to work out the proxy for 4th quarter same store growth.

For computing our own version of these metrics from the cell phone location data, we start with the normalized level of visits (inflated to October 2018 levels), in levels rather than logs. We sum this up for each Starbucks establishment for all three months in the relevant fiscal quarter (which, helpfully, overlap

closely with month ends). We then compute the percentage change from one quarter to the next for each store, winsorized at the 2.5% level in each tail.[12] We then average this over all Starbucks stores to get an average quarterly increase in visits.

This measure imperfectly matches to Starbucks changes in comparable store sales for the quarter in several dimensions. Any variation in how much Starbucks customers are spending at the store will not be captured. The two earnings reports over the sample period where Starbucks discloses both change in transactions and change in sales for the Americas show that these numbers can be considerably different. In the Americas, the fiscal quarter ending April 2017 had sales growth of 3% but transaction growth of -2%, and the quarter ending July 2018 had sales growth of 1% but transaction growth of -2%. Secondly, Starbucks only reports these numbers disaggregated to the segment of the Americas, which includes the US, Canada and Brazil, whereas our data is only for the US. Thirdly, Starbucks rounds its reported sales growth numbers to a whole number of percent. This rounding of +/- 0.5% is very large relative to the variable range, which only goes from 1 to 5 in our sample period. Fourthly, our metrics count all Starbucks locations, whereas Starbucks may treat company-operated and licensed stores differently. Finally, as alluded to earlier, there are a number of different ways that "comparable store sales growth" could be calculated by Starbucks based on particular modeling assumptions and definitions, so we are forced to guess as to what a reasonable metric might involve.

With all these caveats, and given the very small number of quarterly observations, the correlation between our average quarterly change in cellphone visitors and Starbucks reported Americas quarterly comparable store sales growth is 0.85. In Figure 1, we plot the two series next to each other to highlight the visual similarity.

We take these results as supporting the interpretation that our cell phone location data and visitor counts are likely to map strongly to actual Starbucks store-level visits and sales. Most importantly, it

---

[12] Different levels of winsorization in the main cell phone growth measure used in Tables 2 and 3 make very little difference to the results, reinforcing the conclusion that outliers are not driving the differences between Starbucks and other establishments.

provides a consistent metric that can be tracked across both publicly traded and privately-owned business at the establishment level, something very difficult to obtain through other data sources.

Insert Figure 1 here

2. Results

2.1 Starbucks versus Other Establishments

We estimate the effects of Starbucks' bathroom policy using a difference-in-differences framework surrounding the policy enactment (treatment). Our treated group contains all Starbucks stores, whereas the control group includes all non-Starbucks coffee shops and/or other restaurants. Our sample covers January 2017 to October 2018, and since the policy was enacted and publicized in May 2018, we define the treatment period as June 2018 onwards.

Our main hypothesis is that Starbucks with higher exposure will experience greater declines in visits from the public bathroom policy enactment. We conduct our analysis in two steps. First, we study whether there has been a reduction in foot traffic in Starbucks vis-à-vis close-by comparable stores, then we look at the cross section of responses across locations with varying degrees of homeless shelters.

There are two major challenges with this sort of empirical setup. First, the Starbucks announcement could coincide with another event inducing a change in customer preferences for Starbucks and/or control establishments. While our approach accounts for variation in consumer demand for coffee shops through time, we cannot rule out a contemporaneous shock in June 2018 and thereafter – and not the public bathroom policy enactment – that differentially affected Starbucks relative to other coffee shops. We revisit this point when discussing our cross-sectional test, and present an alternative empirical strategy for the sake of robustness, below. However, based on our reading of the media reports around these events, we could not find any publicized event that could potentially create a similar response by the customers around the time of the policy change, other than the arrest of the two men and the associated bathroom policy change. Since our post period only begins in June, we are also measuring changes to monthly visits almost two

16

months *after* the initial period of protests and any negative publicity they may have generated, most of which ended with the announcement of the bathroom policy and the nationwide store closures in May. In this respect, subsequent changes in June and the months afterwards are far more plausibly related to the ongoing effects of the bathroom policy change.

Since we are primarily interested in estimating the effect of Starbucks announcement on the foot traffic of Starbucks vis-à-vis other establishments, we employ the following OLS specification using the sample that includes all establishments classified as being a coffee shop and a sample of non-coffee shop restaurants:

Foot Traffic = $\beta1$ *x Post* + $\beta2$ *x Starbucks$_{it}$ x Post* + $\beta3$ *x Restaurant$_{it}$ x Post*

$$+ \textit{Store Fixed Effects} + \textit{Time Fixed Effects} + \textit{City x Month Fixed Effects} \qquad (1)$$

Our dependent variable, foot traffic, is the natural log of visits to an establishment, observed at a monthly interval. *Post*, our treatment variable, is an indicator taking on a value of one for all months after May 2018. *Starbucks* is an indicator variable which takes a value of one for Starbucks shops. *Restaurant* is an indicator taking on a value of one for establishments classified as being a restaurant, as described in Section 1.1. Depending on the specification, we also include several fixed effects to capture time invariant foot traffic within store, time and city-by-time dimensions. We include *Store Fixed Effects* to absorb unobservable time-invariant characteristics of establishments, including relative differences in general popularity across establishments. However, time-varying effects such as changing local economic conditions could also have an effect on how much consumers spend at retail establishments. This could bias the estimated effect of the treatment if Starbucks establishments are disproportionately located in affected regions. Our specifications therefore include *City x Month* fixed effects (e.g. Dallas *x* July 2018) to account for such differences across cities in each month. Including fixed effects of this nature makes our specification analogous to that recommended by Gormley and Matsa (2014) to control for unobserved heterogeneity.

The coefficient of interest in equation (1) is $\beta 2$ (i.e. *Starbucks * Post)*, as well as the relative size of $\beta 2$ compared to $\beta 3$ (i.e. *Restaurant* Post)*. $\beta 2$ captures how much foot traffic reduced at Starbucks locations following the Starbucks policy enactment, relative to the base case of other coffee shops. $\beta 3$ captures foot traffic at nearby restaurants following the same event compared to the base case of other coffee shops. This is included to partly gauge if a relative difference in foot traffic between Starbucks and other coffee shops is due to a change in Starbucks or because other coffee shops are experiencing unusual increases in relative foot traffic. We use heteroscedasticity-robust standard errors that are double-clustered by CBSA and month.

We present these results in Table 2, which shows that foot traffic declined in Starbucks relative to other coffee shops following the enactment of the bathroom policy. The coefficient on *Starbucks*Post* ranges from -0.054 with only date fixed effects (in column 2), to -0.080 with store and time fixed effects (column 3). In all cases, the estimates are highly significant with *t*-statistics less than -5. A variety of fixed effects specifications are examined – store only (column 1), date only (column 2), store and date (column 3), store and city-by-post (allowing city fixed effects to vary before the policy and after, in column 4), store and city-by-month (as in, separate city effects estimated for each month, in column 5), and store, time and city-by-post (column 6). The decline in visits to Starbucks is large and significant across all specifications.

Insert Table 2 here

The coefficients also represent economically large effects. In our preferred specification in column 5, which includes store fixed effects and city-by-month fixed effects, the coefficient of 0.068 means that Starbucks experienced a 6.8% decline in monthly visits relative to other coffee shops after the enactment of the policy, with a *t*-statistic of -5.94. As discussed before, the inclusion of *Store Fixed Effects* absorbs unobservable time-invariant characteristics of establishments. *City x Month* fixed effects help us control for time-varying effects such as changing local economic conditions which are likely to affect consumer spending across establishments. We also note the absence of any effect for *Restaurant*Post*. Across all

specifications, the coefficients for *Restaurant*Post* are economically small and statistically insignificant, with *t*-statistics less than 0.66. This indicates that other coffee shops appear to resemble nearby restaurants, whereas the changes to Starbucks traffic are strikingly different.

2.2 Store Visits and Distance to Homeless Shelters

Next, we look at whether exposure to Starbucks policy vary across locations. For this purpose, we construct a measure of a given location's exposure to Starbucks' policy. Specifically, we consider the proximity of a given Starbucks store to homeless shelters. We predict that Starbucks that are closer to a homeless shelter would be more likely to attract a group of people that Starbucks customers may not want to interact with. This preference may be due to a host of reasons including associating such homeless people with health deficits and exposure to crime, compared to their non-homeless but impoverished counterparts (Institute of Medicine, 1988, page 39). Being homeless is also associated with shorter life expectancy, higher morbidity and greater usage of acute hospital services (see Hwang et al. (2011) and Kushel et al. (2002). Homeless population also has higher risk for later-stage diagnosis of disease, poor control of manageable conditions (e.g., hypertension, diabetes) and hospitalization for preventable conditions (e.g., skin or respiratory conditions) presumably due to lack of access preventive health services (Rieke et al. (2015). Most importantly, homeless people are the ones most likely to be acutely affected by the policy change. They often lack access to nearby bathrooms, and have difficulty credibly signaling an intent to purchase from a store in order to use their bathrooms. As a result, they seem likely to be especially drawn to the opportunity for free bathrooms and pleasant amenities in which to sit.

For these tests, we consider only establishments located within 20 km of a homeless shelter, so the estimated effects are all measuring only geographic variation within cities that have a homeless shelter in them (rather than differences between the types of cities that do and don't have homeless shelters).

Our base specification takes the following form:

Foot Traffic = β1 *x* Post *x* Distance + β2 *x Starbucks$_{it}$ x* Post *x* Distance

$\qquad$ + β3 *x Restaurant$_{it}$ x* Post *x* Distance + *Store Fixed Effects*

$$+ \textit{ Time Fixed Effects} + \textit{City x Month Fixed Effects} \qquad\qquad (2)$$

We include distance as a continuous variable to estimate the effect of exposure to Starbucks policy. If indeed the presence of the homeless contributes to the decline in visits by other customers, we expect the effect of Starbucks versus other coffee shops to be less and less pronounced in locations that are further away from homeless shelters. We report the results in Table 3 Panel A. In order to aid interpretation, the measure of distance is in kilometers/100. The main variable of importance is *Starbucks*Distance*Post.* This variable is fairly stable across the different fixed effect specifications, and significant at the 5% level in each case. In column 1, with only store fixed effects, it is 0.312 with a *t*-statistic of 2.41, and in the full specification of store and city-by-month fixed effects, it is 0.338 with a *t*-statistic of 2.61.

Insert Table 3 here

In terms of economic magnitude, the base coefficient of *Starbucks*Post* now has the interpretation of the estimated effect right in the vicinity of the homeless shelter. In the full fixed effects version of column 3, this is equal to -0.086, or an 8.6% decline in visits of Starbucks relative to other coffee shops after the policy change. The coefficient on *Starbucks*Distance*Post* of 0.338 means that each additional kilometer of distance from the shelter reduces the size of the Starbucks-vs-other-coffee-shops effect by 0.0034, or 0.34%. For stores 10km from a shelter, the total effect of the policy is thus estimated as -0.086 + 10/100*0.338 = -0.052, or a 5.2% decline in visits.

The tests in Table 3 Panel A all fit an effect of distance based on a continuous linear effect of being further from the homeless shelter. In order to ensure that this is not driving our results, in Panel B we consider alternative specifications for distance. In column 1, we replace the *Starbucks*Post* and *Starbucks*Post*Distance* variables with interactions of four bins for different distances of each store from the homeless shelter: 0-2 km, 2-5 km, 5-10 km, and 10-20 km. Between them, these cover the full range of distances examined, so the four variables of *Starbucks*Post*(Dist<2), Starbucks*Post*(5<Dist<10),* etc.

represent the estimated effect of Starbucks vs other coffee shops at that distance. Column 1 of Panel B presents these results with store and city-by-month fixed effects. The results show that the estimated effect of the policy is monotonic across these categories, being an 8.3% reduction for stores 0-2 km from a shelter, a 7.8% reduction for those 2-5 km away, a 5.3% reduction for those 5-10km away, and a 4.5% reduction for stores 10-20km away.

All of these effects are individually significant at the 1% level or better. However, in this specification the key test is whether the coefficient on *Starbucks*Post*(Dist<2)* is significantly different from the coefficient on *Starbucks*Post*(10<Dist<20)*. This is seen in the F-test at the bottom of the table of 4.58, corresponding to a *p*-value of 0.0458.

In column 2, we perform a similar test to column 1, but instead define breakpoints based on quintiles of distance from the homeless shelter. The results are similar to column 1. Within the closest quintile of distance, Starbucks experienced a decline of 8.2% relative to other coffee shops after the policy change, whereas stores in the furthest quintile experienced a decline of 4.4%. The difference between these two coefficients has a *p*-value of 0.0523. Finally, in column 3 we use the same test as in Panel A, but replace linear distance with the natural log of distance. The effect is still evident, with a coefficient of 0.013 and a *t*-statistic of 2.24. All of these results show that there is an economically large and statistically significant difference between the effect of the policy close to homeless shelters and farther away – the policy had an 84% larger effect for stores less than 2 km from a shelter relative to stores more than 10 km from a shelter.

Next, we compare the effects of homeless shelter distance with one other major component of the urban environment, namely urban density. In particular, we would like to check whether homeless shelter proximity is just measuring the overall property of being in a dense urban area. To this end, we compute two measures of urban density at the zip code level. First one is the density of number of stores (Starbucks, Coffee Shops, and Restaurants). We take the total number of store by month observations and sum it over the whole period at the zip code level, then divide by the land area contained in that zip code. Our second measure is the total number of normalized visits across all stores in the zip code over the whole period scaled by the geographic area in the zip code. Finally, because both measures are highly skewed, we rank

all zip codes as a percentile of this measure, and include this as an interaction term with *Starbucks*Post* and other lower order variables, similar to shelter distance.

These results are presented in Panel C, with all specifications including store and city by time fixed effects. In column 1, we use visitor count density interactions, and find a significant negative effect of the interaction of *Starbucks * Post * Visitor Density*. Because density ranges from zero to one as a percentile measure, the coefficient of -0.093 (with a *t*-statistic of -2.23) means that the densest zip code sees a 9.3% reduction in visitors relative to the least dense (the *Starbucks * Post* coefficient, here an insignificant 1.2% increase). The effect of store density in column 2 is somewhat weaker, with a coefficient of -0.065 (*t*-statistic of -1.68). In columns 3 and 4, we repeat the same two regressions, but also include homeless shelter distance as an interaction term as well. The coefficient on *Starbucks * Post * Distance* is now 0.343 (*t*-statistic of 2.66) and 0.344 (*t*-statistic of 2.67) respectively after controlling for visitor and store density, respectively. This is very similar to the equivalent coefficient in Panel A column 3 of 0.338, meaning that controlling for urban density makes almost no difference to the estimated effect of distance from a homeless shelter. By contrast, the coefficients on density are reduced about a third for visitor density (-0.065), and sixty percent for store density (-0.027), with neither now statistically significant. This indicates that a reasonable amount, but not all, of the estimated effect of urban density is actually measuring the effect of proximity to a homeless shelter.

As noted earlier, it is difficult to fully disentangle the effects of homeless shelters mattering directly, versus homeless shelters being a proxy for other aspects of the urban environment. Part of the challenge is that knowing the precise location of homeless populations is, by its very nature, quite difficult. There are good theoretical reasons to predict that homeless populations will be particularly affected by the policy change, and the effects seem to be robust to proxies for general urban density. However, with these results, we are unable to rule out the alternative possibility that some other aspect of geography that is correlated with homeless shelter location is driving a differential effect between Starbucks and other coffee shops after the policy enactment, and not the presence of the homeless themselves.

2.3. Alternative Method: Synthetic Control

Recall the difference-in-difference model employed in the preceding analysis is accompanied by the identifying assumption that, absent the policy intervention, treated stores (Starbucks) would not differ from non-treated stores (e.g., other coffee shops) after May 2018. This assumption might be violated if Starbucks experienced a relative decline in popularity relative to other coffee shops around the same time that the bathroom policy was enacted. In contrast, while our fixed effects do not provide perfect identification, it becomes more difficult to explain the differential effects for stores near homeless shelters with a similar identification concern. More precisely, the gap cannot be driven by Starbucks as a whole getting better or worse for reasons other than the policy (such as offering new products, better service, etc.), as this is common at all distances. It cannot be driven by worsening economic conditions around the homeless shelter, or even differences in preferences for coffee around the homeless shelter. Alternative theories would need to explain why Starbucks got worse relative to other coffee shops by a larger amount for stores closer to homeless shelters. Nonetheless, in this section, we briefly consider possible challenges to the results of Table 3.

Figure 2 illustrates the difference in store traffic for Starbucks compared to other coffee shops as a function of time and the distance of an establishment to a homeless shelter. Specifically, we turn to the four distance bins described in the first specification of Table 3 Panel A above. Figure 2 reports the difference in the mean of logged visits to Starbucks establishments relative to other coffee shops for each month and distance bin. Two stylized facts emerge. First, in the months leading up to the enactment of the policy there does not appear to be a systematic divergence in the relative popularity of Starbucks relative to other coffee shops across the different distance bins, inconsistent with potential challenges to the conclusions drawn from Table 3. Second, the difference in relative performance of Starbucks appears to slightly decline relative to other coffee shops in the months leading up to the policy change (reversing a positive spike in the several months around the end of 2017). While this may be due to many factors, such as seasonality in Starbucks popularity, it raises potential questions regarding the conclusions drawn from the preceding

analysis regarding the effect of the policy across different geographic regions. For this reason, we seek additional validation of the results presented to this point before continuing.

Insert Figure 2 here

To account for the potential pre-treatment deviation of foot-traffic to Starbucks relative to other coffee shops, we turn to the synthetic control method pioneered by Abadie and Gardeazabal (2003) and Abadie, Diamond, and Hainmueller (2012). Intuitively, rather than rely on all non-Starbucks coffee shops to serve as the control group, this method constructs a synthetic control observation for each treated observation by forming a convex combination of non-treated observations (non-Starbucks coffee shops) that most closely resembles the treated observation in the pre-treatment period. While there are many dimensions over which one may attempt to maximize the similarity between the synthetic control observations and treated observations, a natural choice is the outcome variable (logged store visits) in the months prior to policy intervention.[13] More precisely, for each Starbucks store, we construct a convex combination of other coffee shops that minimizes the difference in logged store visits between the treated and synthetic control observation in the full time-series prior to the policy change. For tractability, we restrict the sample of candidate observations to those non-Starbucks coffee shops residing in the same 3-digit ZIP code.[14] Next, we re-consider the analysis performed in Table 3 under this alternative framework.

We consider the estimated effects of the policy as a function of an establishment's distance to a homeless shelter in Table 4. Following the change in methodology, the estimated effect continues to be more pronounced in Starbucks locations near homeless shelters. Moreover, the absolute difference in point estimates between stores in the closest distance bin and those in the farthest bin closely resemble the estimates in Table 3. Indeed, the relative importance of close-to-shelters stores versus far-from-shelters

---

[13] In fact, this is the precise example described in Imbens and Wooldridge (2009).

[14] In instances in which the pool of candidate controls is less than 200 observations in size, we extend the pool to include neighboring ZIP codes until the pool is sufficiently large to exceed this threshold.

stores is much larger under synthetic controls. The overall reduction is now almost exclusively driven by the nearby stores, so the percentage difference (which was 84% in Table 3) now becomes enormous. Moreover, far-away stores under synthetic controls now show now statistically insignificant decreases in estimated foot-traffic. Taken together, these results continue to support the conclusion that the bathroom policy resulted in a decrease in Starbucks store visits, and that the effect is being driven by stores near a homeless shelter.

Insert Table 4 here

2.4. Demographics of Customer Changes

In Tables 5 and 6, we investigate whether customer response to the Starbucks policy differs across demographic characteristics. In particular, we take two measures obtained from census block level demographic information on visitors to each Starbucks – the average income level of customers, and the percentage of customers who are white. We take these store-level aggregated measures of census-block average income and percentage of white residents, and run similar tests to those in Table 2, to find out if Starbucks experienced different demographic changes relative to other coffee shops.

The income measure relates to the hypothesis that wealthier customers may have different preferences against associating with homeless populations. This could arise because they have less exposure to the poor in their day to day lives. The race measure relates to the fact that the original controversy centered in party around allegations that the enforcement of the previous customers-only bathroom policy was done in a racist manner.[15] The publicity from the policy change may have had the effect of highlighting perceptions that Starbucks was previously acting in a racist manner, or, conversely, to greater appreciation for the change in policy as being more racially enlightened. Such effects, however arising, may be expected

---

[15] See, for instance, https://money.cnn.com/2018/04/19/news/companies/starbucks-arrests-philadelphia/index.html?iid=EL. Relatedly, Starbucks closed their stores for an afternoon so employees could undergo training in understanding racial biases.

to differentially impact black and white customers. By contrast, once the bathroom policy itself was implemented, the direct effects observed so far mostly relate to homeless populations, where the predictions for how different racial groups will respond are less clear.

The results for income are presented in Table 5. In this table, the dependent variable is average income of customers. We find that the average income of Starbucks customers decreased relative to the average income of other coffee shops over the treatment period. The coefficients are all -0.003 regardless of the fixed effects used, with t-statistics between -1.83 and -2.52. This suggests that the decline in Starbucks attendance documented in previous tables was greater among its wealthier clientele. This would be consistent with them being more sensitive to crowding and the new visitors brought in by the bathroom policy. In this specification, restaurants in general also saw a significant decline in average income relative to the control group of other coffee shops. As a consequence, it is harder to rule out the possibility that some change in other coffee shops is driving the trend. Nonetheless, to the extent that coffee shops and restaurants do not behave in an identical fashion in terms of customer demographics at all points in time, the most relevant comparison is the decline in income among Starbucks visitors relative to other coffee shops.

Insert Table 5 here

In Table 6, we do the same tests as above, when replacing average income with the percentage of white residents from the census block level. Notably, we do not find any significant differences in the effect of Starbucks policy on its racial customer composition relative to other coffee shops. The point estimates are all directionally positive but very small (less than or equal to 0.1 percentage point changes in the fraction of white customers) and statistically insignificant. The result of course does not speak to either Starbucks motivation for the original policy, nor its actual effect on different racial groups. Rather, it is informative of the attitudes of different races of customers towards the change in bathroom policy that ultimately resulted from the protests. If there is no difference in the racial customer composition after the policy, then white Starbucks customers appear equally bothered by the new policy as non-white Starbucks customers.

26

2.5.   Time Spent in Store

While the results reported so far inform us on the extensive margin of the treatment, e.g. the number of customer visits to an establishment, they are silent about the intensive margin, i.e. how much visiting customers spent in an establishment. Because we do not have access to store level revenues, we rely on a measure that our cell phone data can provide, minutes spent in the store. As we discussed in the data section, devices send periodic pings to service providers with time between pings ranging from a few seconds to several minutes. By using the time between pings and the location of the pings, we can estimate the number of minutes a given device spent at a given location. We replace our *Foot Traffic* metric with the natural log of the estimated dwell time, the *Minutes Spent in Store* metric, to estimate whether remaining customers spent less or more time per store visit following the policy enactment.

The results are presented in Table 7. In the full specification with store and city-by-month fixed effects, customers spend on average 4.2% less time in Starbucks relative to other coffee shops after the policy change, with a *t*-statistic of 5.28. In this test, *Restaurants* show an increase relative other coffee shops, meaning that there was a relative decline in the length of visits for coffee shops in general relative to restaurants, but a significantly larger decline for Starbucks relative to other coffee shops. In other specifications, we continue to find similar results which suggest not only fewer customers visited Starbucks stores but remaining visitors spent less time in the store. This is consistent with greater table utilization and bathroom use resulting in people not lingering in a Starbucks store, or leaving without making a purchase. In this sense, the small number of non-paying visitors who *do* linger and use tables and bathrooms have an outsized effect on the total number of visitors, who either stop coming and/or spend less time in the store.

Because we do not have information on how much customers spent on average when they visit a given store, we cannot speak to overall revenue impact of policy enactment. However, given the reduction in both extensive and intensive margins, it is unlikely that moving from the quasi-public bathroom policy

to a completely open public bathroom has benefited Starbucks unless the customers increased their purchase significantly to make up for the intensive and extensive margin declines. The fact that Starbucks visitors also came disproportionately from lower income areas after the policy also militates against this possibility, as the increase in purchases would have to be driven by Starbucks' relatively poorer segment of its previous customer base. In addition, none of this considers any extra staffing costs involved in greater bathroom maintenance.

Insert Table 7 here

In our final test, we change our focus from establishments to events occurring around the establishments. Specifically, we investigate whether the new policy increased or decreased forms of crime that can be plausibly related to the amenities offered at Starbucks locations. For this purpose, we collect incident-level micro-data reported by three cities (i.e. Austin, Denver and Pittsburgh) on public urination related crimes. The idea here is that people urinating and defecating on the streets has clear negative externalities for the people in the surrounding areas, and an establishment that offers facilities to visitors regardless of whether they perform a transaction essentially provides a solution to a public health problem. To test this, we look at census block groups in these cities before and after the Starbucks policy change, and compare their distance to a Starbucks. The main variable of interest is *Post*Distance_to_Starbucks*. In this analysis, our dependent variable is the number of citations for urination-related crimes. We include census block group fixed effects (to account for the fact that areas have differential citation counts in general) and either month fixed effects or city-by-month fixed effects, to strip out general time-series changes in crime rates.

These results are presented in Table 8 Panel A. The coefficient on *Post*Distance_to_Starbucks* is 0.206 with a *t*-statistic of 2.15 with census-block and city-by-month fixed effects, in column 2. When distance is measured as the log of distance in columns 3 and 4, the effect is stronger, with the full fixed effects specification giving a coefficient of 0.296 with a *t*-statistic of 2.42. In terms of economic magnitude

in column 2, a block group that is one standard deviation further from a Starbucks (1.19 km) will have public urination citations in the post period that are higher by 24.5% (0.206 * 1.19 = 0.245) relative to the closer block group, so the effect is economically large.

These patterns may simply reflect generally changing levels of crimes in these areas, however, and not public urination specifically. To test this, in Panel B we conduct the same full specification tests (columns 2 and 4) for a range of placebo citations in the same three cities and time period. These are disturbing the peace, simple assaults/fighting, marijuana possession, shoplifting, theft of service, threats/harassment, and vandalism. As well as being of similar criminal severity, these crimes are chosen based on all three cities reporting non-trivial numbers of geo-coded crimes with a sufficiently similar description to be able to be classified.[16] All numbers of citations are again scaled by the block group average, so the coefficients are of comparable magnitude across the different crime types.

Panel B indicates that the geographical effect of Starbucks on public urination citations in the post period is not present for any of the other crimes examined. Regardless of whether distance is measured in logs or levels, none of the other crimes show any statistically significant effects for the *Post * Distance to Starbucks*. In addition, the magnitude of the coefficients is considerably smaller - the closest is around half as large, and more than half the coefficients are negative in sign.

These results are consistent with a bathroom channel for our main results, but difficult to explain otherwise. People most likely to urinate in public are those most likely to be affected by the new policy, and the change in this variable as a function of distance to Starbucks lends support for bathrooms as part of the mechanism for the change in sales. Moreover, this result is difficult to explain by other competing explanations. For instance, suppose one believed that Starbucks was somehow becoming worse in a way that differentially affected customers near homeless shelters, patron income, and time spent in stores, but *not* due to the new policy. In this case, it is not clear why this should be related to time-by-geography

---

[16] Other similar kinds of offenses are excluded based on not being present in all cities. For instance, Austin and Pittsburgh have citations for public drunkenness, but Pittsburgh only lists a general "other public order crimes" category. Pittsburgh and Denver report crimes for liquor possession, Austin does not, etc.

changes in public urination crimes. Moreover, it is not clear why public urination should show such a different pattern to a range of other similar minor crimes relating to disorderly behavior.

Insert Table 8 here

3. Conclusion

A fundamental prediction of public economics is that when an investment has a personal cost but a common benefit, individuals will underinvest. Because of this free rider problem, the private market often under-supplies public goods. The private sector can, in some cases, try to overcome the free rider problem of providing public goods by charging user fees that are proportional to their valuation of the public good. The Starbucks experiment provides an opportunity to observe the effects on customer behavior of providing a particular form of public good (bathrooms) at zero cost to everyone. Our evidence suggest that the Starbucks policy decreased both the foot traffic and time spent at Starbucks, especially in those that are closer to homeless shelters.

Our results highlight the difficulty of policies that try to shift public goods provision onto private companies. While certain public goods can be bundled with sales so that it is still profitable for companies to make them available, at a certain point stores must decide how much to curtail the provision of the good to non-patrons, who may also actively deter others' purchases. The negative consequences of Starbucks policy suggest that profit-maximizing companies will be likely to underprovide exactly for those groups who have the lowest willingness to pay, in this case due to extreme poverty. To the extent that publicly available bathrooms are viewed as an important public good, our results suggest that their provision ought to be the business of the government, rather than relying on private corporations.

References

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller, 2012. Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program, *Journal of the American Statistical Association*, 105 (490), 493-505.

Abadie, Alberto and Javier Gardeazabal, 2003. The Economic Costs of Conflict: A Case Study of the Basque Country, *The American Economic Review*, 93 (1), 113-132.

Andreoni, James, 1990. Impure altruism and donations to public goods: A theory of warm-glow giving. *Economic Journal* 100.401 (1990): 464-477.

Athey, Susan, Billy Ferguson, Matthew Gentzkow, and Tobias Schmidt, 2019. Experienced Segregation, Stanford University Working Paper.

Bagnoli, Mark, and Susan G. Watts, 2003. Selling to socially responsible consumers: Competition and the private provision of public goods. *Journal of Economics and Management Strategy* 12.3 (2003): 419-445.

Baron, David P., 2007, Corporate social responsibility and social entrepreneurship. *Journal of Economics and Management Strategy* 16.3: 683-717.

Bergstrom, Theodore, Lawrence Blume, and Hal Varian, 1986. On the private provision of public goods. *Journal of Public Economics* 29.1: 25-49.

Bolton, Gary E., and Elena Katok, 1998. An Experimental Test of the Crowding-Out Hypothesis: The Nature of Beneficent Behavior. *Journal of Economic Behavior and Organization* 37 (3): 315–31.

Chava, Sudheer, 2014. Environmental externalities and cost of capital. *Management Science* 60.9: 2223-2247.

Long, Elisa F., M. Keith Chen, and Ryne Rohla, 2019. Political Storms: Tracking Hurricane Evacuation Behavior using Smartphone Data. UCLA Working Paper.

Chen, M. Keith and Ryne Rohla, 2018. The effect of partisanship and political advertising on close family ties.", *Science* 360, 1020–1024.

Dimson, Elroy, Oguzhan Karakas, and Xi Li, 2010. Active ownership. *Review of Financial Studies* 28.12: 3225-3268.

Edmans, Alex. 2011. Does the stock market fully value intangibles? Employee satisfaction and equity prices. *Journal of Financial Economics* 101.3, 621-640.

Ferrell, Allen, Hao Liang, and Luc Renneboog. 2016. Socially responsible firms. *Journal of Financial Economics* 122.3: 585-606.

Fisman, Raymond, Shachar Kariv, and Daniel Markovits. 2007. Individual Preferences for Giving. *American Economic Review* 97 (5): 1858–76.

Friedman Milton, 1970. The social responsibility of business is to increase its profits. New York Times Magazine (September 13), https://www .nytimes.com/1970/09/13/archives/article-15-no-title.html

Institute of Medicine (US). Committee on Health Care for Homeless People, 1988. *Homelessness, health, and human needs*. National Academies.

Hartzmark, Samuel M. and Abigail B. Sussman. 2019, Do Investors Value Sustainability? A Natural Experiment Examining Ranking and Fund Flows, *Journal of Finance* (Forthcoming), 2019.

Hwang, Stephen W., et al., 2011, Health status, quality of life, residential stability, substance use, and health care utilization among adults applying to a supportive housing program. *Journal of Urban Health* 88.6, 1076-1090.

Imbens, G.W. and Wooldridge, J.M., 2009. Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, *47*(1), pp.5-86.

Kruger, Philipp, 2015. Corporate goodness and shareholder wealth. *Journal of financial Economics* 115.2, 304-329.

Kushel, Margot B., et al., 2002, Emergency department use among the homeless and marginally housed: results from a community-based study. *American Journal of Public Health* 92.5: 778-784.

Morgan, John, and Justin Tumlinson, 2019, Corporate provision of public goods, *Management Science* .

Servaes, Henri, and Ane Tamayo, 2013, The impact of corporate social responsibility on firm value: The role of customer awareness. *Management Science* 59.5, : 1045-1061.

Ottoni-Wilhelm, Mark, Lise Vesterlund, and Huan Xie. *2017*. Why Do People Give? Testing Pure and Impure Altruism. *American Economic Review* 107 (11): 3617-33.

Ribar, David C., and Mark O. Wilhelm. 2002. Altruistic and Joy-of-Giving Motivations in Charitable Behavior. *Journal of Political Economy* 110 (2): 425–57.

Rieke, Katherine, et al.. 2015, Mental and nonmental health hospital admissions among chronically homeless adults before and after supportive housing placement. *Social Work in Public Health* 30.6, 496-503.

Yildirim, Huseyin, 2014, Andreoni-McGuire Algorithm and the Limits of Warm-Glow Giving, *Journal of Public Economics* 114: 101–107.

Figure 1 – Growth in Starbucks Cell Phone Visits and Reported Same-Store Growth
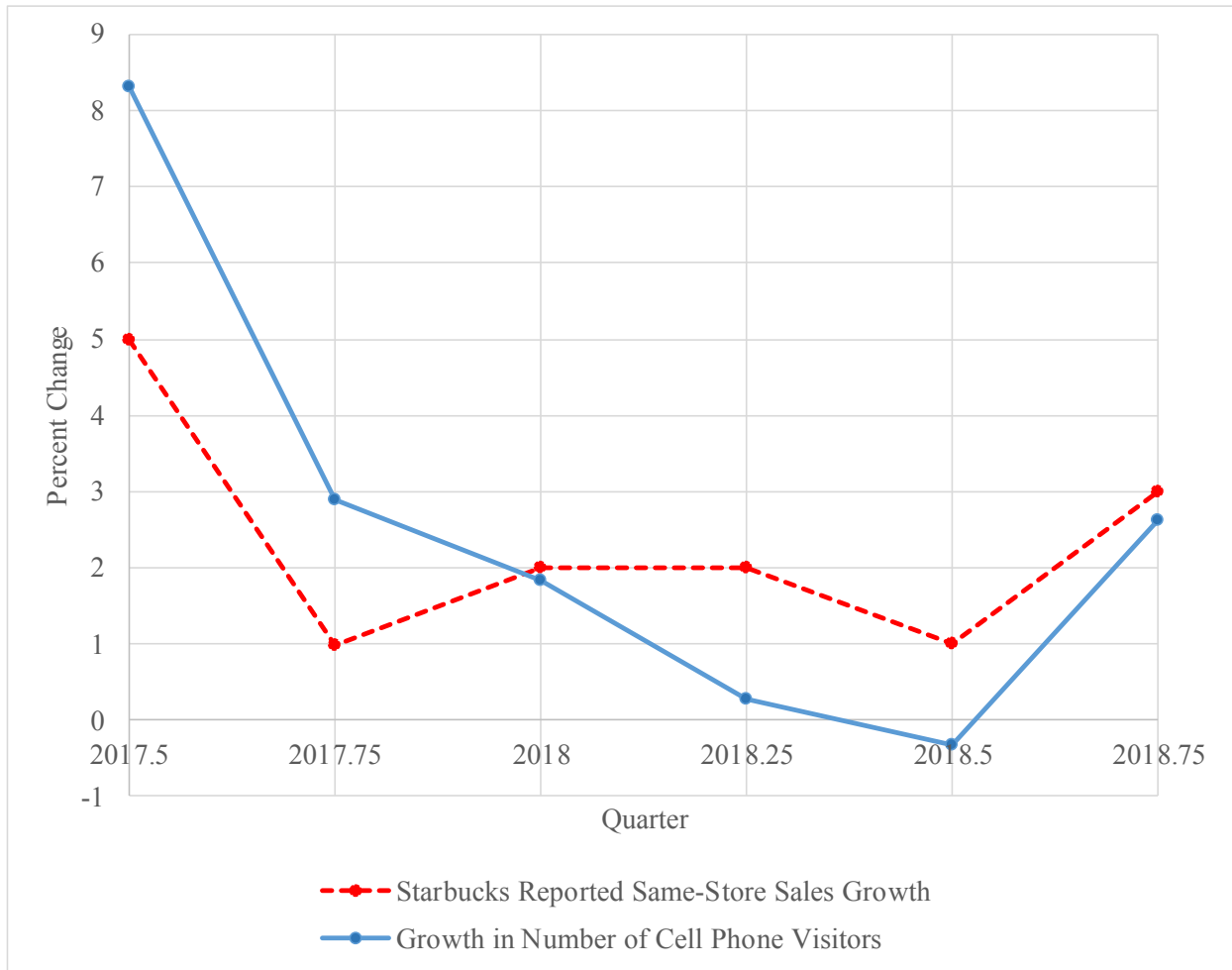
This figure plots changes in cell phone visits and publicly reported measures of increased patronage of Starbucks stores. The dashed line is taken from Starbucks quarterly and annual reports, and is the average same-store sales growth for Starbucks stores in the Americas (with fourth fiscal quarter numbers, being those ending in ".75", estimated from annual and quarterly numbers). The solid blue line is the average percentage increase in the normalized number of visitors to Starbucks establishments in the United States based off the main cell phone location data.

Figure 2 – Trends in Starbucks Versus Other Coffee Shops Before Policy Change, Split by Distance
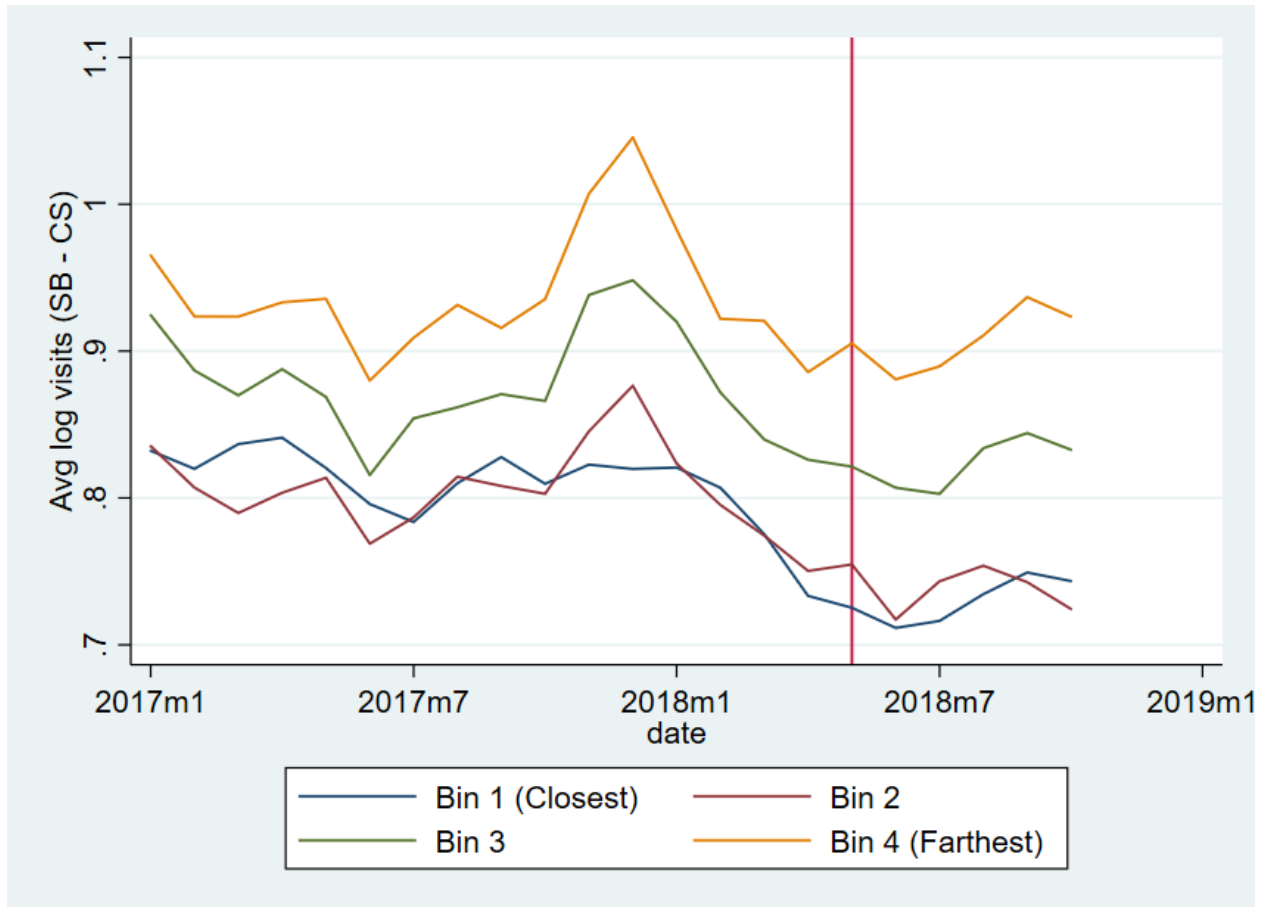
This figure plots the difference in average log normalized visits between Starbucks and other coffee shops, split by the distance from a homeless shelter. We begin with the main dependent variable from Tables 2 and 3 – the log of establishment visits, normalized based on the city-wide growth in device usage to October 2018 levels. This is then averaged by month, store category (Starbucks versus Coffee Shops) and binned distance from homeless shelter (0-2 km, 2-5 km, 5-10 km, 10-20 km). We then compute the difference between Starbucks and other Coffee shops for each month/bin combination, and plot it in the above graph. The red vertical line is the first post-treatment date, namely June 2018.

Table 1. Descriptive Statistics

This table reports summary statistics for the sample considered. In Panels A and B, US Cell phone location data runs from January 2017 to October 2018. Visits are based on anonymized cell phone location pings being within the store's footprint, and are aggregated up to the monthly level. "Starbucks" refers to Starbucks stores, "Coffee Shops" refers to all other coffee shops, and "Restaurants" is a 25% random sample of remaining restaurants. "Distance to Shelter" is calculated only for establishments in cities that have a homeless shelter. Panel C gives census-block level measures of crimes relating to public urination, from Austin, Denver and Pittsburgh.

Panel A: All Stores						
	N	Mean	SD	P25	P50	P75
Number Visits	3,678,722	344	473	102	220	427
Dwell Time (mins)	3,681,209	35.1	58.6	13.0	24.0	42.0
Distance to Shelter (km)	3,363,234	7.67	8.84	2.02	4.76	10.22
Est. Income ($1k)	2,277,891	70.9	31.0	49.3	65.1	87.3

Panel B: By Store Type						
	N	Mean	SD	P25	P50	P75
Starbucks:						
Number Stores	10,752					
Number Visits	231,279	675	717	280	512	834
Dwell Time (mins)	231,410	22.8	50.3	9.0	14.0	21.0
Distance to Shelter (km)	224,448	6.96	7.17	2.23	4.87	9.52
Est. Income ($1k)	192,186	82.3	32.6	59.8	77.5	100.3
Coffee Shops:						
Number Stores	24,800					
Number Visits	511,238	318	520	97	207	374
Dwell Time (mins)	511,477	30.5	57.9	9.0	19.0	36.0
Distance to Shelter (km)	474,859	7.12	8.65	1.65	4.24	9.38
Est. Income ($1k)	307,190	73.6	32.3	51.1	67.9	91.1
Restaurants:						
Number Stores	137,745					
Number Visits	2,936,205	322	428	98	209	404
Dwell Time (mins)	2,938,322	36.8	59.1	15.0	26.5	43.5
Distance to Shelter (km)	2,663,927	7.83	8.99	2.07	4.84	10.43
Est. Income ($1k)	1,778,515	69.3	30.2	48.2	63.4	84.9

Panel C: Incident-Level Sample						
	N	Mean	SD	P25	P50	P75
Distance to Shelter (km)	12,600	1.50	1.19	0.63	1.17	1.97
Number of Incidents	12,600	0.09	0.68	0.00	0.00	0.00

Table 2. Starbucks Visits Versus Other Similar Establishments After Bathroom Policy Change

This table reports the results of OLS regressions where the dependent variable is the natural log of visits to an establishment, observed at a monthly interval between January 2017 and October 2018. The sample consists of Starbucks stores, other coffee shops and sample of non-coffee shop restaurants. *Starbucks* is an indicator variable taking on a value of one for a Starbuck establishment. *Restaurant* is an indicator taking on a value of one for establishments classified as being a restaurant (with other coffee-shops being the omitted category). *Post* is an indicator taking on a value of one for all months after May 2018, when Starbucks implemented its change in bathroom policy. Reported *t*-statistics in parentheses are heteroscedasticity-robust and double-clustered by CBSA and month.

	Dependent Variable is Log of Monthly Store Visits					
Starbucks * Post	-0.080***	-0.054***	-0.080***	-0.067***	-0.068***	-0.068***
	(-6.10)	(-6.35)	(-6.15)	(-5.90)	(-5.94)	(-5.27)
Restaurant * Post	0.004	0.010	0.004	0.006	0.006	
	(0.24)	(0.66)	(0.23)	(0.30)	(0.30)	
Post	0.101***					
	(6.20)					
Starbucks		0.898***				
		(28.29)				
Restaurant		0.051*				
		(2.07)				
Store Fixed Effects	Yes	No	Yes	Yes	Yes	Yes
Time Fixed Effects	No	Yes	Yes	No	No	Yes
City by Post Fixed Effects	No	No	No	Yes	No	Yes
City by Time Fixed Effects	No	No	No	No	Yes	No
Observations	3,694,153	3,694,981	3,694,153	3,694,153	3,694,110	3,694,149
R-squared	0.916	0.035	0.918	0.916	0.920	0.919

Table 3. Starbucks Visits and Distance to Homeless Shelters After Bathroom Policy Change

This table reports the results of OLS regressions where the dependent variable is the natural log of visits to an establishment, observed at a monthly interval between January 2017 and October 2018. The sample consists of Starbucks stores, other coffee shops and sample of non-coffee shop restaurants, for stores within 20km of a homeless shelter. *Post* is an indicator taking on a value of one for all months after May 2018, when Starbucks implemented its change in bathroom policy. In Panel A, *Distance* is defined as the Euclidean distance, in kilometers, of an establishment to the nearest homeless shelter. In Panel B, distance is variously measured as bins for 0-2 km, 2-5 km, 5-10 km and 10-20 km (column 1), quintiles of distance (column 2) or log of distance (column 3). In Panel C, we include interactions of the main effects with two measures of urban density, *Visitor Density* and *Store Density*, into the specification. To compute *Store Density*, we take the total number of store by month observations and sum it over the whole period at the zip code level, then divide by the land area contained in that zip code. We compute *Visitor Density* as the total number of normalized visits across all stores in the zip code over the whole period scaled by the geographic area in the zip code. Because both measures are highly skewed, we rank all zip codes as a percentile of *Store Density* and *Visitor Density*. Where not absorbed by fixed effects, all lower order interaction variables are also included in the regression. All remaining variables are defined in Table 2. Reported *t*-statistics in parentheses are heteroscedasticity-robust and double-clustered by CBSA and month.

	Panel A - Baseline Regressions of Shelter Distance		
	Dependent Variable is Log of Monthly		
Starbucks * Post * Distance	0.312**	0.310**	0.338**
	(2.41)	(2.40)	(2.61)
Restaurant * Post * Distance	-0.274***	-0.275***	-0.247***
	(-3.70)	(-3.70)	(-3.22)
Starbucks * Post	-0.094***	-0.094***	-0.086***
	(-7.01)	(-7.05)	(-6.57)
Restaurant * Post	0.136	0.141	0.012
	(1.48)	(1.51)	(0.12)
Post * Distance	0.020	0.020	0.022
	(1.22)	(1.22)	(1.24)
Post	0.083***		
	(5.18)		
Store Fixed Effects	Yes	Yes	Yes
Time Fixed Effects	No	Yes	No
City by Time Fixed Effects	No	No	Yes
Observations	3,092,918	3,092,918	3,092,906
R-squared	0.920	0.922	0.923

	Panel B - Alternative Versions of Homeless Shelter Distance Regressions		
	Dependent Variable is Log of Monthly Store Visits		
Starbucks * Post * (Dist < 2)	-0.083***		
	(-5.52)		
Starbucks * Post * (2 < Dist < 5)	-0.078***		
	(-6.43)		
Starbucks * Post * (5 < Dist < 10)	-0.053***		
	(-4.07)		
Starbucks * Post * (10 < Dist < 20)	-0.045**		
	(-2.66)		
Starbucks * Post * (Dist q1)		-0.082***	
		(-5.44)	
Starbucks * Post * (Dist q2)		-0.088***	
		(-5.99)	
Starbucks * Post * (Dist q3)		-0.070***	
		(-5.60)	
Starbucks * Post * (Dist q4)		-0.052***	
		(-3.77)	
Starbucks * Post * (Dist q5)		-0.044**	
		(-2.67)	
Starbucks * Post * Log Distance			0.013**
			(2.24)
Store Fixed Effects	Yes	Yes	Yes
City by Time Fixed Effects	Yes	Yes	Yes
F-test for [SB*Post*Close] = [SB*Post*Far]	4.58	4.23	N/A
p-value	0.0443	0.0523	N/A
Observations	3,092,906	3,092,906	3,092,906
R-squared	0.923	0.923	0.923

	Panel C - Urban Density and Shelter Distance			
	Dependent Variable is Log of Monthly Store Visits			
Starbucks * Post	0.012	-0.013	-0.035	-0.063*
	(0.33)	(-0.41)	(-1.04)	(-1.98)
Starbucks * Post * Distance			0.343**	0.344**
			(2.66)	(2.67)
Starbucks * Post * Visitor Density	-0.093**		-0.059	
	(-2.23)		(-1.47)	
Starbucks * Post * Store Density		-0.065		-0.027
		(-1.68)		(-0.74)
Store Fixed Effects	Yes	Yes	Yes	Yes
Time Fixed Effects	No	No	No	No
City by Time Fixed Effects	Yes	Yes	Yes	Yes
Observations	3,078,138	3,078,138	3,078,138	3,078,138
R-squared	0.923	0.923	0.923	0.923

Table 4. Robustness for Distance Results using Synthetic Control Method

This table reports the results of OLS regressions where the dependent variable is the natural log of visits to an establishment, observed at a monthly interval between January 2017 and October 2018. The table repeats the analysis performed in Table 3 using the synthetic control method. Accordingly, the sample of non-Starbucks coffee shops from the previous analysis is replaced with a synthetic control group best able to match the pre-treatment outcome variables for each of the treated observations. For each treated observation, we restrict the sample of candidate non-treated observations to coffee shops within the same 3-digit ZIP code. *Post* is an indicator taking on a value of one for all months after May 2018, when Starbucks implemented its change in bathroom policy. Reported *t*-statistics in parentheses are heteroscedasticity-robust and double-clustered by CBSA and month.

	Dependent Variable is Log of Monthly Store Visits			
Starbucks * Post * Linear Distance	0.414***			
	(4.46)			
Starbucks * Post * Log Distance		0.017***		
		(4.17)		
Starbucks * Post * (Dist < 2)			-0.041***	
			(-5.04)	
Starbucks * Post * (2 < Dist < 5)			-0.027***	
			(-3.33)	
Starbucks * Post * (5 < Dist < 10)			-0.006	
			(-0.68)	
Starbucks * Post * (10 < Dist < 20)			0.009	
			(0.78)	
Starbucks * Post * (Dist q1)				-0.041***
				(-4.74)
Starbucks * Post * (Dist q2)				-0.036***
				(-4.49)
Starbucks * Post * (Dist q3)				-0.018*
				(-1.88)
Starbucks * Post * (Dist q4)				-0.005
				(-0.48)
Starbucks * Post * (Dist q5)				0.008
				(0.76)
Store Fixed Effects	Yes	Yes	Yes	Yes
City by Time Fixed Effects	Yes	Yes	Yes	Yes
F-test for [SB*Post*Close] = [SB*Post*Far]	N/A	N/A	16.42	14.98
p-value	N/A	N/A	0.0006	0.0009
Observations	2,856,684	2,856,684	2,856,684	2,856,684
R-squared	0.922	0.922	0.922	0.922

Table 5. Income of Customers After Bathroom Policy Change

This table reports the results of OLS regressions where the dependent variable is the natural log of estimated income for customers visiting an establishment, observed at a monthly interval. Income estimates are based on the census block group average income from the 2017 Census Bureau American Community Survey, weighted by the visits per residential census block group. The sample consists of Starbucks stores, other coffee shops and sample of non-coffee shop restaurants. All remaining variables are defined in Table 2. Reported *t*-statistics in parentheses are heteroscedasticity-robust and double-clustered by CBSA and month.

	Dependent Variable is Log Customer Zip Code Income		
Starbucks * Post	-0.003*	-0.003*	-0.003**
	(-2.01)	(-1.83)	(-2.52)
Restaurant * Post	-0.005***	-0.006***	-0.005***
	(-3.70)	(-3.75)	(-4.38)
Post	0.008**		
	(2.81)		
Time Fixed Effects	No	Yes	No
Store Fixed Effects	Yes	Yes	Yes
City by Time Fixed Effects	No	No	Yes
Observations	2,275,118	2,275,118	2,275,023
R-squared	0.855	0.855	0.857

Table 6. Race of Customers After Bathroom Policy Change

This table reports the results of OLS regressions where the dependent variable is the estimated percentage of customers who are white visiting an establishment, observed at a monthly interval. Customer race estimates are based on the census block group percentage of white residents from the 2017 Census Bureau American Community Survey, weighted by the visits per residential census block group. The sample consists of Starbucks stores, other coffee shops and sample of non-coffee shop restaurants. All remaining variables are defined in Table 2. Reported *t*-statistics in parentheses are heteroscedasticity-robust and double-clustered by CBSA and month.

	Dependent Variable is Customer Zip Code Pct White		
Starbucks * Post	0.000	0.001	0.001
	(0.77)	(1.31)	(1.25)
Restaurant * Post	-0.001	-0.001*	-0.001*
	(-1.70)	(-1.73)	(-1.98)
Post	-0.003***		
	(-2.96)		
Time Fixed Effects	No	Yes	No
Store Fixed Effects	Yes	Yes	Yes
City by Time Fixed Effects	No	No	Yes
Observations	2,284,333	2,284,333	2,284,241
R-squared	0.941	0.941	0.942

Table 7. Time Spent in Store After Bathroom Policy Change

This table reports the results of OLS regressions where the dependent variable is the natural log of the minutes spent in the establishment, observed at a monthly interval. The sample consists of Starbucks stores, other coffee shops and sample of non-coffee shop restaurants. All remaining variables are defined in Table 2. Reported *t*-statistics in parentheses are heteroscedasticity-robust and double-clustered by CBSA and month.

	Dependent Variable is Avg Minutes Spent in Store		
Starbucks * Post	-0.043***	-0.042***	-0.042***
	(-4.11)	(-4.02)	(-5.28)
Restaurant * Post	0.033***	0.033***	0.040***
	(6.34)	(6.23)	(6.82)
Post	-0.027		
	(-1.20)		
Time Fixed Effects	No	Yes	No
Store Fixed Effects	Yes	Yes	Yes
City by Time Fixed Effects	No	No	Yes
Observations	3,694,153	3,694,153	3,694,110
R-squared	0.783	0.785	0.789

Table 8. Crime Related to Public Urination After Bathroom Policy Change

This table reports the results of OLS regressions to examine how the scaled number of crimes of each particular type in a geographical Census block group changed after Starbucks bathroom policy as a function of distance from a Starbucks store. Each value is the monthly number of citations for that crime type scaled by the average number of crimes in the block group over the full sample. Crime data is taken for Austin, Denver and Pittsburgh (which collect geo-coded crime incident locations), from January 2016 to December 2018. Distance to Starbucks represents the distance, in kilometers or log of kilometers respectively, from the block group's centroid to the nearest Starbucks. All remaining variables are defined in Table 2. In Panel A, the crime in question is citations for public urination. In Panel B, a range of other crime types are considered – disturbing the peace, simple assault/fighting, marijuana possession, shoplifting, theft of service, threats and harassment, and vandalism. These are chosen based on having non-trivial numbers of classifiable crimes of similar description in all three cities. Reported t-statistics in parentheses are heteroscedasticity-robust and double-clustered by block group and month.

	Panel A - Public Urination Citations			
	Dependent Variable is Scaled Citations for Urination-Related Crimes			
Post * Distance to Starbucks	0.188**	0.206**		
	(2.05)	(2.15)		
Post * Log Distance to Starbucks			0.271**	0.296**
			(2.38)	(2.42)
Census Block Group Fixed Effects	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	No	Yes	No
City by Time Fixed Effects	No	Yes	No	Yes
Observations	12,600	12,600	12,600	12,600
R-squared	0.003	0.011	0.003	0.011

	Public Urination	Disturbing the Peace	Assaults, Fighting	Marijuana Possession	Shoplifting	Theft of Service	Threats and Harrassment	Vandalism
				Panel B - Public Urination vs Other Crimes				
Post * Distance to Starbucks	0.206**	0.015	0.042	-0.117	-0.098	0.043	0.105	-0.019
	(2.15)	(0.25)	(0.87)	(-1.66)	(-0.91)	(0.22)	(1.37)	(-0.42)
Census Group Block Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City by Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,600	11,520	10,656	8,856	7,344	6,336	11,592	10,260
R-squared	0.011	0.017	0.012	0.020	0.019	0.022	0.018	0.019

	Public Urination	Disturbing the Peace	Assaults, Fighting	Marijuana Possession	Shoplifting	Theft of Service	Threats and Harrassment	Vandalism
Post * Log Distance to Starbucks	0.296**	-0.016	0.015	-0.121	-0.106	-0.038	0.142	-0.021
	(2.42)	(-0.18)	(0.15)	(-1.03)	(-0.75)	(-0.29)	(1.47)	(-0.20)
Census Group Block Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City by Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,600	11,520	10,656	8,856	7,344	6,336	11,592	10,260
R-squared	0.011	0.017	0.012	0.020	0.019	0.022	0.018	0.019