JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS Vol. 58, No. 4, June 2023, pp. 1473–1507

doi:10.1017/S0022109022001442

Measuring and Improving Stakeholder Welfare Is Easier Said than Done

Umit G. Gurun
University of Texas at Dallas Department of Finance
umit.gurun@utdallas.edu

Jordan Nickerson
University of Washington Department of Finance
jnick@uw.edu

David H. Solomon
Boston College Department of Finance
david.solomon@bc.edu (corresponding author)

Abstract

While corporate social responsibility by firms aims at improving welfare for different social groups, whether it achieves this is often difficult to measure. After Apr. 2018 protests, Starbucks enacted policies that anybody could sit in their stores and use the bathroom without making a purchase. Using anonymized cellphone location data, we estimate this led to a 7.0% decline in attendance relative to other nearby coffee shops. The effect is 84% larger near homeless shelters and larger for Starbucks' wealthier customers. The average time spent per visit declined by 4.1%. Public urination citations decreased near Starbucks locations, but other minor crimes were unchanged.

I. Introduction

Recent years have witnessed an increased debate over the extent to which firms should engage in corporate social responsibility (CSR), with companies providing altruistic or prosocial products, services, or practices. One school of thought pioneered by Friedman (1970) argues managers should solely focus on profit maximization within legal constraints. Standing in opposition is the view that firms should focus on doing good for other stakeholders, but that doing so can also increase shareholder value (the "Strategic CSR" view in Vishwanathan, van

An earlier draft of this article was titled "The Perils of Private Provision of Public Goods." We are grateful to Auren Hoffman, Lauren Spiegel, and Noah Yonack for their help with providing and understanding the SafeGraph's anonymized GPS data. We also thank Sam Hartzmark, Eugene Soltes, and seminar participants at Chapman University, the University of St. Gallen, the 2021 Western Finance Association Meetings, and the 2021 European Finance Association Meetings for helpful comments. Some of the results in this article are based on researchers' own analyses calculated (or derived) based in part on data from The Nielsen Company (US), LLC and marketing databases provided through the Nielsen data sets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the Nielsen data are those of the researchers and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein. All remaining errors are our own.

Oosterhout, Heugens, Duranc, and van Essen (2020) and Bénabou and Tirole (2010)). However, the literature on "doing well by doing good" mostly focuses on evaluating the first half of "doing well," and generally just takes the second part of "doing good" as given. That is, the main study is whether firm shareholders benefit from CSR actions, and it is mostly just assumed that nonshareholder groups are actually made better off.

However, for this result to be true, two additional pieces are required. First, managers must be able to observe the values that different stakeholders place on possible policies, even before they can decide how to aggregate these into an overall measure of welfare. This task is made more complicated by the fact that some stakeholders do not generate observable feedback in the form of market signals, so even determining what stakeholders jointly want is not an easy task. Second, the marginal spending on CSR must in fact increase net social benefit, rather than simply evincing good intentions. The potential for adverse or unintended consequences expands further when managers choose to make CSR an integral part of business operations, rather than simply transferring cash to a prosocial cause. In such cases, CSR runs the risk of undermining the profit-generating core of the business, with flow-on effects on a firm's customers, suppliers, and other such groups that also comprise a social welfare calculation. These two aspects (measuring stakeholder preferences and actually improving social welfare) operate in addition to the third piece of whether firm shareholders themselves benefit.

To this end, we study the difficulty of satisfying these three claims of CSR in the context of Starbuck's recent decision to provide public amenities to noninvestor stakeholders. On Apr. 15, 2018, a Starbucks store in Philadelphia called the police after 2 African–American men refused to leave the store, despite not purchasing anything. This led to a series of nationwide protests accusing Starbuck's existing policies of exhibiting racial bias (https://nbcnews.to/36GXVTI). In response, Starbucks held a day of sensitivity training for all employees on May 29, 2018, 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.[1]

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 use the bathroom as long as one looked like they *might* be about to make a purchase (or officially use it for the price of a purchase).[2] To move from a mostly open 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 clienteles, then the change could nonetheless have significant effects.

We explore this question using anonymized cellphone location data from more than 10 million devices from Jan. 2017 to Oct. 2018. We estimate monthly visits to

[1]The half-day of training and closure of stores nationwide was announced on Apr. 17, 2018. The policy change was reported in *The Wall Street Journal* on May 19, 2018. See https://bit.ly/30N1bJz and https://on.wsj.com/34FUXvY.

[2]This mirrors remarks made by Howard Schulz around the time of the incident (https://cnn.it/3uktRWh).

each Starbucks location, for roughly 74% of Starbucks' US stores where GPS data can be measured reliably. Our empirical approach compares the change in demand for Starbucks relative to other nearby coffee shops (e.g., Peet's, Coffee Bean & Tea Leaf), and other local restaurants. Our baseline specification suggests that Starbucks stores experienced a 7.0% decrease in visits after the enactment of the policy, compared with similar coffee shops and restaurants. This gap is consistent across various specifications that control for different average levels of store visits, time trends, and city-month fixed effects to ensure we are not just measuring differences in local economic conditions. Consistent with the effects we document, Starbucks experienced large negative market-adjusted stock returns of −11.12% on June 20 and 21, 2018, when they released negative earnings guidance for the second quarter. Other coffee-related stocks went up 0.58% over the same 2 days.

Strikingly, the decrease in visits after the policy enactment is significantly larger for locations closer to homeless shelters. Stores less than 2 km away experienced declines of 8.5% relative to nearby coffee shops, while stores more than 10 km away experienced declines of only 4.8%. 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, and increased use by the general public of bathrooms and tables is also estimated to have negative impacts. The relative decline in Starbucks visits is also greater in denser urban environments, consistent with foot traffic also creating more demand for bathrooms. However, controlling for population density leaves the homeless shelter effect largely unchanged. Because it is difficult to know the location of the homeless populations, proximity to homeless shelters may be proxying for other aspects of the urban environment that increase the effect of the policy, not just the effect of the homeless themselves.

Additional supportive evidence for a negative effect of the policy comes from the duration of customer visits. 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 4.1%, again relative to other coffee shops and restaurants. If the negative effects were merely reputation, customers would seem more likely to avoid the store altogether, rather than turn up briefly and then leave. Using weekly household-level coffee purchase information gathered from a large panel of households, we find that after the policy change, households from zip codes with a greater share of Starbucks visits prior to the policy change saw significantly greater increases in retail (i.e., noncoffee shop) home coffee purchases relative to those zip codes where Starbucks was less popular. This suggests a substitution away from in-store Starbucks purchases toward more home coffee consumption.

Starbucks also experienced a significant change in the demographics of those 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 0.4%, 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

https://doi.org/10.1017/S0022109022001442 Published online by Cambridge University Press

of 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 public urination citations near Starbucks locations relative to other areas after the policy change. By contrast, a wide range of other minor public order crimes shows no significant changes or consistent signs of effects. This result is especially difficult to explain 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 bathroom policy had a direct effect that was 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. 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. This view contrasts one claim of stakeholder capitalism as leading to an increase in firm profitability – rather, profitability allows Starbucks to engage in CSR and bear the associated loss because of the surplus generated by other activities. Our assumption is that other coffee shops are the counterfactual for how Starbucks sales would have changed absent the new bathroom policy. It is possible that the counterfactual of no new bathroom policy would have resulted in a greater revenue loss, through bad publicity, further incidents, lawsuits, and so forth. Nonetheless, the large difference between effects close to and far from homeless shelters (around half of the baseline effect), suggests a considerable component due to the bathroom policy itself, even if inaction would have had a larger overall cost.

This episode also highlights the tradeoffs firms face when deciding whether to provide public amenities, and the potential impact on net social benefits. As a store broadens the provision of amenities from customers, to potential customers, to people unlikely to be customers, and to noncustomers who may actually *deter* other customers, scarce store resources get consumed with less and less private return in response. The cost can actually be lost sales and decreased consumer surplus, not just greater staffing costs to keep the bathrooms clean. This illustrates the difficulty of applying stakeholder theories of management. While Starbucks' shareholders may disagree on what will improve firm value, their theoretical interests are fairly well aligned. But once shareholder value is not the decisive metric, how to trade-off the benefits of wider bathroom usage versus costs to existing customers and employees is much less obvious.

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. Both are examples of consumers' total utility being driven by more than the items directly purchased, though the implications are slightly different. 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 whom 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.[3]

Our article contributes to the literature on whether firms ought to increase their CSR activities. As Karpoff (2021) notes, theories of CSR are most compelling when they explain or justify managers deviating from simple NPV-maximizing rules. Arguably, CSR actions that lead to increased shareholder value ought to be undertaken even in a neoclassical setting (Hart and Zingales (2017), Pástor, Stambaugh, and Taylor (2021)). Under the Friedman (1970) view that advocates pure profit maximization (subject to legal constraints), CSR may provide cover for managers to waste corporate resources for personal gain. This "shareholder interests" view does not require firms to be solely exploitative; as Friedman notes, "firms can generate profits *only* by developing strong and mutually beneficial relationships with customers, suppliers, and employees" (Karpoff (2021)). But CSR activities beyond this are likely to be value-destroying (e.g., Cheng, Hong, and Shue (2013)). This prediction would also hold if CSR has large positive externalities and managers aim to serve a broader social purpose, but these effects are not internalized in shareholder value (Matten, Crane, and Chapple (2003)).

In contrast, a long literature claims to find a panoply of benefits associated with CSR for firms, which would seem to predict that Starbucks ought to have benefited from the new policy. These include increased innovation and resilience,[4] a lower cost of capital and greater access to finance,[5] lower risk,[6] better performance,[7] better customer and employee relationships,[8] attracting customers and talented employees,[9] and as a form of advertising.[10] However, in the literature that argues that CSR improves corporate performance, it is often unclear if this relation is meant to only hold locally for endogenously chosen CSR (which might represent actions consistent with the simple NPV-maximization rule), or meant to hold in a global sense (e.g., Bénabou and Tirole (2010), Liang and Renneboog (2020), and Vishwanathan et. al. (2020)). In other words, is the trade-off of additional CSR a net positive for all firms or for every possible amount of additional CSR? While few

[3]The two explanations are not entirely distinct, if customers feel that the ambiance of the store is affected by the presence of the homeless, and this is considered one of the amenities that customers consume.

[4]See Luo and Du (2015) and Ortiz-de-Mandojana and Bansal (2016).

[5]See Dhaliwal, Li, Tsang, and Yang (2011), El Ghoul, Guedhami, Kwok, and Mishra (2011), and Cheng, Ioannou, and Serafeim (2014).

[6]See Kim, Park, and Weir (2012), Koh, Qian, and Wang (2014), and Sun and Cui (2014).

[7]See Mackey, Mackey, and Barney (2007), Edmans (2011), Eccles, Ioannou, and Serafeim (2014), Flammer (2015), and Lins, Servaes, and Tamayo (2017).

[8]See Greening and Turban (2000), Jones, Willness, and Madey (2014), and Bode, Singh, and Rogan (2015).

[9]See Becker-Olsen, Cudmore, and Hill (2006), Luo and Bhattacharya, (2006), Bhattacharya, Sen, and Korschun (2008), Brekke and Nyborg (2008), Gregg, Grout, Ratcliffe, Smith, and Windmeijer (2011), and Liang and Renneboog (2017).

[10]Servaes and Tamayo (2013). For other benefits, see also Krüger (2015), Masulis and Reza (2015), Ferrell, Liang, and Renneboog (2016), Lins, Servaes, and Tamayo (2017), and Dai, Liang, and Ng (2021).

authors argue this explicitly, the strong implication in such papers tends to be "… and therefore managers ought to do more CSR," and it is often difficult to know when CSR is expected to *not* increase shareholder value. While understanding these effects is complicated because companies rarely engage in prosocial behavior by chance, the nature of Starbuck's policy change provides one such instance to evaluate the firm profitability claim of stakeholder capitalism.

Most importantly, our article fills a gap in the literature by studying the impact of CSR on the stakeholder groups affected by it, a question that has received much less attention. While Starbucks' bathroom policy was likely intended as a strategic CSR investment, our results suggest that it had a direct negative effect on sales. Importantly, our article also evaluates the effect on nonshareholder groups. Our results demonstrate a large negative impact on customers, and an implied cost to suppliers from reduced sales, but positive effects for the relatively small number of people directly affected by the bathroom policy, and from less public urination. In doing so, our article highlights both the necessity and difficulty of considering the effects on shareholders and nonshareholders alike in the welfare calculation, underscoring the challenges of efficiently engaging in CSR.

II. Data and Sample Construction

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

A. Establishment-Level Foot Traffic

The establishment-level foot traffic is provided by SafeGraph, a company that aggregates anonymized smartphone-location data from numerous smartphone apps (e.g., local news and weather) 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, the number of pings, and 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 5 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 Jan. 2017 to Oct. 2018. Finally, our data allows us to estimate the demographics of visitors to an establishment in a given month, such as race, income, and so forth. We infer these traits for each visitor by matching the census block group the device resides into the income and racial share for the block group from the 2017 American Community

https://doi.org/10.1017/S0022109022001442 Published online by Cambridge University Press

Survey from the Census Bureau. For each store months visits, we compute the weighted-average income and race of visitors from these residential block group demographic shares.

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 that 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 722,515 (SNACK_AND_NONALCOHOLIC_BEVERAGE_BARS). From this set, we hand-classify each firm with at least 5 store locations based on the company name (and web search if necessary) to determine its eligibility for the COFFEE_SHOP group. As our second criterion, we consider an establishment to belong to COFFEE_SHOP if the i) firm's name contains the word "coffee," and ii) 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 72,251, 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, it is 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) (https://pewrsr.ch/3ckQhQK). 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 smartphone 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 detrend our foot traffic data. More precisely, we first define TOTAL_COUNT_{ct} as the total number of visits made by all devices in core-based statistical area (CBSA) c in month t.[11] We then scale the number of visits to establishment i in CBSA c in month t by $\text{TOTAL\_COUNT}_{ct}/$

[11]A CBSA is a U.S. geographic area that combines 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.

https://doi.org/10.1017/S0022109022001442 Published online by Cambridge University Press

$TOTAL\_COUNT_{cT}$, where T is the final month in our sample. Thus, each visit count is adjusted to an Oct. 2018 level based on the overall growth in that city.[12]

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

B. Homeless Shelter Locations

Our second data set contains homeless shelter locations and addresses collected from two sources: the homeless shelter directory (www.homelessshelterdirectory.org) and the Google Places API. This data set 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 *U.S. 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.

C. Incident-Level Crime Reports

We collect the incident-level microdata from 2016 to 2018 reported by 3 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 3 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 aggregate up instances of public urination at a monthly interval and 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 with at least one instance of public urination over the period considered, yielding a final sample of 350 block groups.

[12]The main results of the article are similar in magnitude and significance of raw visit counts (i.e., uncorrected for growth in the number of cellphones) are used instead.

D. Financial Data

Finally, we take daily stock returns for Starbucks and other publicly traded coffee-related companies from the Center for Research in Security Prices. We take analyst information from IBES.

E. 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 345 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 smartphone application from which our data provider obtains location data. For stores in cities that have a homeless shelter recorded in our data set, the average store is located 7.18 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 $71 k. Panel B partitions the sample based on the 3 categories we study. While COFFEE_SHOPS and RESTAURANTS are similar in their number of visitors

TABLE 1

Descriptive Statistics

Table 1 reports summary statistics for the sample considered. In Panels A and B, US Cell phone location data runs from Jan. 2017 to Oct. 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.

	N	Mean	Std. Dev.	P25	P50	P75
Panel A. All Stores						
No. of visits	3,246,388	345	476	101	219	428
Dwell time (mins)	3,681,209	35.1	58.6	13.0	24.0	42.0
Distance to shelter (km)	3,366,269	7.18	8.59	1.76	4.26	9.49
Est. income ($1 k)	2,277,891	70.9	31.0	49.3	65.1	87.3
Panel B. By Store Type						
Starbucks						
No. of stores	10,706					
No. of visits	193,721	688	740	282	518	848
Dwell time (mins)	231,410	22.8	50.3	9.0	14.0	21.0
Distance to shelter (km)	224,625	6.48	6.98	1.94	4.38	8.79
Est. income ($1 k)	192,186	82.3	32.6	59.8	77.5	100.3
Coffee shops						
No. of stores	24,045					
No. of visits	455,278	318	530	97	207	374
Dwell time (mins)	511,477	30.5	57.9	9.0	19.0	36.0
Distance to shelter (km)	475,537	6.66	8.40	1.42	3.77	8.73
Est. income ($1 k)	307,190	73.6	32.3	51.1	67.9	91.1
Restaurants						
No. of stores	137,846					
No. of 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 ($1 k)	1,778,515	69.3	30.2	48.2	63.4	84.9
Panel C. Incident-Level Sample						
Distance to shelter (km)	12,600	1.50	1.19	0.63	1.17	1.97
No. of incidents	12,600	0.09	0.68	0.00	0.00	0.00

https://doi.org/10.1017/S0022109022001442 Published online by Cambridge University Press

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 the number of stores, we have nonmissing visit data for 10,706 Starbucks locations, 24,045 coffee shops, and 137,846 restaurants. This is less than the 14,620 Starbucks locations that SafeGraph has business listing and footprint data for.[13] 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 are stand-alone locations.

Finally, Panel C of Table 1 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.

F. 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, and so forth. More usefully, Starbucks also reports its own measure of percentage change in comparable store sales, rounded to a whole number of percent. Discussions with Starbucks' Investor Relations department indicate that Starbucks computes this growth for a quarter relative to same-quarter sales in the previous year. 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 the fourth fiscal quarter specifically. This makes interpreting fourth quarter revenues fairly straightforward, as the difference between the total

[13]Starbucks 10-K from Nov. 2018 lists 14,606 US stores, comprising 6,031 licensed stores and 8,575 company-operated stores, as of Sept. 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 United States.

and the previous 3 quarters. Fourth quarter same-store growth numbers can only be approximated, however. We approximate them based on annual and quarterly growth numbers, plus quarterly sales revenues, as described in the Appendix.

For computing our own version of these metrics from the cell phone location data, we start with the normalized level of visits (inflated to Oct. 2018 levels), in levels rather than logs. We sum this up for each Starbucks establishment for all 3 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.[14] 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 Apr. 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, we compute growth from one quarter to the next, rather than using the same quarter from a year prior. To do the latter with our limited time series of cell phone data would only give us three observations of quarterly growth to compare, and also places a larger strain on the assumption that the city-level correction for overall device count increases is being fully controlled for (as consecutive quarters will have less impact from such inflation relative to quarters a year apart).[15]

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

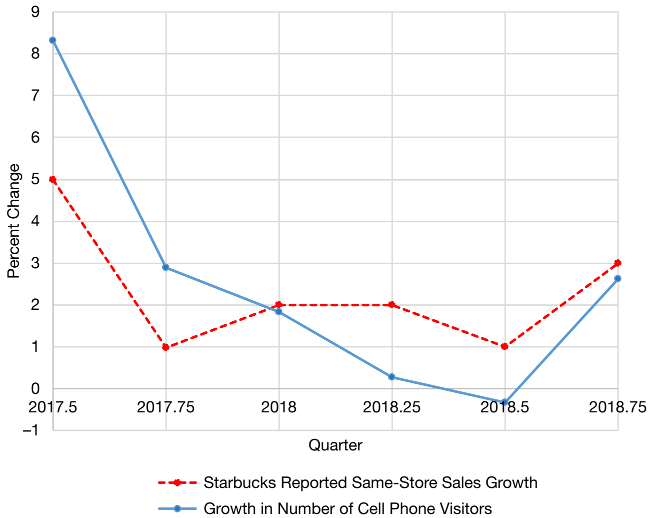
[14]Different levels of winsorization in the main cell phone growth measure used in Tables 2 and 3 make very little difference to the results, reinforcing the conclusion that outliers are not driving the differences between Starbucks and other establishments.

[15]The 3 year-over-year estimates we can compute for Starbucks stores are for Q1, Q2, and Q3 of 2018, and are a 17.6% increase, an 8.9% increase, and a 7.8% increase, respectively. These are all considerably higher than Starbucks' comparable store sales growth numbers over the same period (increases of 2%, 1%, and 3%, respectively), but also less than the numbers for coffee shops over the same period (23.2% increase, 17.4%, and 15.2% increases, respectively). This seems to suggest that trying to apply the exact same methodology at long horizons with the cell phone data may induce more noise than it fixes. In other words, if the largest artificial variation from true growth is due to seasonal variation, year-over-year estimates will give a more accurate picture of growth, whereas if the largest artificial variation is an overall secular trend, closer-in-time estimates will be less distorted.

FIGURE 1

Growth in Starbucks Cell Phone Visits and Reported Same-Store Growth

Figure 1 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 "0.75," 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 on the main cell phone location data.



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

III. Results

A. 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 Jan. 2017 to Oct. 2018, and since the policy was enacted and publicized in May 2018, we define the treatment period as June 2018 onward.

Our main hypothesis is that Starbucks establishments 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 distance to homeless shelters.

There are two major challenges with this 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 customer response around the time of the policy change, other than the arrest of the two men and the associated bathroom policy change. Importantly, since our post-period only begins in June 2018, we are measuring changes to monthly visits almost two 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 afterward are more plausibly related to the ongoing effects of the bathroom policy change, and not the initial protests and publicity.

The maintained assumption from this setup is that, absent the new bathroom policy, Starbucks' changes in visits after June 2018 would have resembled those of other coffee shops. Because our main dependent variable is the log of foot traffic, our effects approximately measure changes month to month in foot traffic, and our treatment variable thus measures changes beginning in June. To the extent that media interest in public scandals tends to be rather short-lived before moving on to the next scandal, it seems likely that after 6 weeks (when the bathroom policy was enacted) Starbucks was not going to continue to get big increases in bad publicity or customer reactions if it took no further actions. One possible alternative is if the lack of policy change might lead to *further* scandals – in other words, if Starbucks decided that it was untenable to maintain a policy of store manager discretion without having similar incidents and that each new one would be worse due to the cumulative effect. Even in such a hypothetical, Starbucks still had a choice in which way managerial discretion was removed – they equally could have enacted a strict "customers only" bathroom and tables policy, which would not have had the effects we document. Nonetheless, if such counterfactual ongoing problems would have been present, our estimates will document the direct cost of the policy, but not the net (possible) savings from avoiding further incidents. An alternate counterfactual is that absent the policy change some share of Starbucks' original customer base would seek out alternate business in light of the scandal. While we cannot observe this alternative, for Starbucks' policy to represent an optimal response, this counterfactual would need to result in an even larger drop in foot traffic. To identify bathroom channels specifically, we rely more heavily on distance-to-homeless-shelter tests, which are less subject to this critique.

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

$$\begin{aligned}\text{FOOT\_TRAFFIC} = {} & \beta 1 \times \text{POST} + \beta 2 \times \text{STARBUCKS} \times \text{POST} \\ & + \beta 3 \times \text{RESTAURANT} \times \text{POST} \\ & + \text{STORE\_FIXED\_EFFECTS} + \text{TIME\_FIXED\_EFFECTS} \\ & + \text{CITY} \times \text{MONTH\_FIXED\_EFFECTS}.\end{aligned}$$

(1)

Our dependent variable, FOOT_TRAFFIC, is the natural log of visits to an establishment, observed at a monthly interval. POST is an indicator taking on a value of 1 for all months after May 2018. STARBUCKS is an indicator variable that takes a value of 1 for Starbucks shops, with the uninteracted version of this variable being omitted due to the presence of store-fixed effects. RESTAURANT is an indicator equal to 1 for establishments classified as a restaurant, as described in Section II.A. 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 × MONTH fixed effects (e.g., Dallas × 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 was 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 after the enactment of the bathroom policy. The coefficient on STARBUCKS × POST ranges from −0.049 with only date-fixed effects (column 2), to −0.073 with store and time-fixed effects (column 3). In all cases, the estimates are highly significant with *t*-statistics of approximately −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,

TABLE 2

Starbucks Visits Versus Other Similar Establishments After Bathroom Policy Change

Table 2 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 Jan. 2017 and Oct. 2018. The sample consists of Starbucks stores, other coffee shops, and a random sample of noncoffee shop restaurants. STARBUCKS is an indicator variable taking on a value of 1 for a Starbucks establishment. RESTAURANT is an indicator taking on a value of 1 for establishments classified as being a restaurant (with other coffee-shops being the omitted category). POST is an indicator taking on a value of 1 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. *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

	Dependent Variable: log(MONTHLY_STORE_VISITS)					
STARBUCKS × POST	−0.073***	−0.049***	−0.073***	−0.070***	−0.070***	−0.073***
	(−5.69)	(−4.91)	(−5.71)	(−5.77)	(−5.79)	(−5.42)
RESTAURANT × POST	0.004	0.007	0.004	0.003	0.003	
	(0.22)	(0.51)	(0.22)	(0.13)	(0.13)	
POST	0.046***					
	(2.99)					
STARBUCKS		0.891***				
		(23.05)				
RESTAURANT		0.036				
		(1.07)				
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
No. of obs.	3,245,648	3,246,388	3,245,648	3,245,648	3,245,627	3,245,644
R^2	0.917	0.033	0.917	0.917	0.918	0.917

time and city-by-post (column 6). The decline in visits to Starbucks is large and significant across all specifications.

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 −0.070 means that Starbucks experienced a 7.0% decline in monthly visits relative to other coffee shops after the enactment of the policy, with a *t*-statistic of −5.79. As discussed before, the inclusion of STORE_FIXED_EFFECTS absorbs unobservable time-invariant characteristics of establishments. CITY × 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.

Another estimate of economic magnitude is to turn the coefficients into estimated dollar costs. We take Starbucks 2017 10-K, which reports more detailed breakdowns of revenues within the Americas. We take Starbucks' 2017 fiscal year revenues for the Americas in the Company-operated and Licensed Stores ($15.613b), multiply it by our baseline reduction (7.0%), and multiply by the 5 out of 12 months we estimate the effect for. Finally, to account for the fact that the Americas includes stores in Canada and Latin America, we multiply by the ratio of SafeGraph Starbucks locations in the US (14,620) to the total stores in the Americas in 2017 (16,559). This gives us a total estimated reduction in sales from the policy of \$402 m over the months we measure. If we conservatively only apply the reduction to the 10,706 Starbucks stores for which we have cell phone data, and assume no reduction elsewhere, this corresponds to a \$294 m decrease in sales.

One aspect worth noting is that the difference-in-difference framework we utilize assumes that Starbucks would have followed the path of other coffee shops but for the policy change. However, if there is a significant substitution between Starbucks and other coffee shops after the change (e.g., Starbucks patrons decide to go to a different coffee shop instead), then the effects will be overstated. This substitution has a flavor of a cross-price elasticity of demand in microeconomics, except that the variables changing here are not price, but rather the provision of complements to the goods actually being purchased. Conceptually, it seems unlikely that a substitution effect would result in more than one visitor arriving at other coffee shops for each patron that switches from Starbucks. In this sense, one might expect a reasonable lower bound on the magnitudes from substitution to be a division by two in the coefficient. The main specification only concretely estimates a difference between the two types of stores from the policy, not a single-store-type policy effect.

Nonetheless, there is some suggestive evidence that the majority of the effect is coming from Starbucks itself, as seen in the small and insignificant coefficient on RESTAURANT $\times$ POST. If other coffee shops were showing large growth (such as under a substitution explanation), one might expect visitor growth to outpace nearby restaurants, who will be largely unaffected. This does not appear to be going on, as other coffee shops show visitor growth that resembles restaurants. However, interpreting this is not straightforward, as anything else that drives overall coffee demand will also affect the differential trend between coffee shops and restaurants, so the lack of an effect, while being consistent with most of the effect being driven by Starbucks itself, is not strong evidence either way.

Even if other coffee shops are not being affected due to the policy change, it remains possible that Starbucks itself would have suffered worse consequences in bad publicity, future negative incidents and lawsuits, lost sales, and so forth, if they *had not* implemented something like the current policy. This counterfactual is necessarily difficult to test cleanly. In this version of events, the overall effect of the new policy would have a direct impact of the bathroom change on customers' preferences, and a second publicity effect of redeeming some of Starbucks' lost reputation. It is possible that the overall effect of these two channels is positive, notwithstanding that the change in visits after the policy is negative. The direct costs of the incident (in terms of avoiding future repetitions) are hard to ascertain, as the legal settlement between the men and Starbucks was confidential. As one benchmark, the men announced a settlement with the city of Philadelphia for \$1 each, plus a \$200,000 fund to help entrepreneurs (https://nyti.ms/3gd7Bby). It seems likely that the situation resembles of costs to firms of financial misrepresentation as described in Karpoff, Lee, and Martin (2008), where the indirect reputational penalties are much larger than the direct legal penalties. This could include lower demand, negative political responses against the firm, and other potential problems. In this version of events, the negative costs of the bathroom policy would be a form of tax that Starbucks has to pay to overcome its reputation of racism.

This version, while possible and hard to rule out, is not the inevitable interpretation, however. The alternative perspective is that these kinds of negative media events are powerful but short-lived, and attention would have eventually moved elsewhere. The post-period in our sample begins in June 2018, which is after the initial negative publicity of the incident.

Our interest, however, is less in identifying the Starbucks-specific reputation effect, and more with identifying the bathroom channel itself. This aspect speaks to the more general problem of public amenities and the implications of CSR on customer behavior. To better identify this aspect, we turn to additional tests that are strong predictions of where a bathroom channel ought to show larger effects, but which are less obviously predicted by lingering effects of reputational costs.

B. Store Visits and Distance to Homeless Shelters

To identify a bathroom channel more tightly, we examine the effects of the policy across locations according to how likely they were to be negatively affected by bathroom effects. 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 nonhomeless but impoverished counterparts (Institute of Medicine (1988), p. 39). Being homeless is also associated with shorter life expectancy, higher morbidity, and greater usage of acute hospital services (see Kushel, Perry, Bangsberg, Clark, and Ross (2002), and Hwang, Gogosis, Chambers, Dunn, Hoch, and Aubry (2011)). The homeless population also has a 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 to preventive health services (Rieke, Smolsky, Bock, Erkes, Porterfield, and Watanabe-Galloway (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.[16] As a result, they seem likely to be especially drawn to the opportunity for free bathrooms and pleasant amenities.

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

[16]It is also worth noting that many homeless shelters primarily provide housing in the evenings, but are closed during the day. In this sense, being near a homeless shelter does not necessarily correspond to being near a bathroom for most of the hours in which Starbucks stores are open.

https://doi.org/10.1017/S0022109022001442 Published online by Cambridge University Press

between the types of cities that do and do not have homeless shelters). Our base specification takes the following form:

$$\begin{aligned}(2)\quad \text{FOOT\_TRAFFIC} = {} & \beta 1 \times \text{POST} \times \text{DISTANCE} + \beta 2 \times \text{STARBUCKS} \\ & \times \text{POST} \times \text{DISTANCE} + \beta 3 \times \text{RESTAURANT} \\ & \times \text{POST} \times \text{DISTANCE} + \text{STORE\_FIXED\_EFFECTS} \\ & + \text{TIME\_FIXED\_EFFECTS} + \text{CITY} \\ & \times \text{MONTH\_FIXED\_EFFECTS}.\end{aligned}$$

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 Panel A of Table 3. 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 0.330 with a *t*-statistic of 2.29, and in the full specification of store and city-by-month fixed effects, it is 0.310 with a *t*-statistic of 2.16.

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 −0.088, or an 8.8% decline in visits to Starbucks relative to other coffee shops after the policy change. The coefficient on STARBUCKS × DISTANCE × POST of 0.310 means that each additional kilometer of distance from the shelter reduces the size of the Starbucks-vs-other-coffee-shops effect by 0.0031, or 0.31%. For stores 10 km from a shelter, the total effect of the policy is thus estimated as $-0.088 + 10/100 \times 0.31 = -0.057$, or a 5.7% decline in visits.

The tests in Panel A of Table 3 all model the impact of distance based on a continuous linear effect of being further from the homeless shelter. 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× DISTANCE and POST × DISTANCE variables with interactions of 4 bins for different distances of each store from the homeless shelter: 0–2, 2–5, 5–10, and 10–20 km. These cover the full range of distances examined, so the 4 variables of STARBUCKS × POST × (DIST < 2), STARBUCKS × POST × (5 < DIST < 10), and so forth represent the estimated effect of Starbucks versus other coffee shops at that distance. Column 1 in Panel B presents these results with store and city-by-month fixed effects. The estimated effect of the policy is monotonic across these categories, being an 8.5% reduction for stores 0–2 km from a shelter, a 7.7% reduction for those 2–5 km away, a 6.5% reduction for those 5–10 km away, and a 4.8% reduction for stores 10–20 km away.

TABLE 3

Starbucks Visits and Distance to Homeless Shelters After Bathroom Policy Change

Table 3 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 Jan. 2017 and Oct. 2018. The sample consists of Starbucks stores, other coffee shops, and a random sample of noncoffee shop restaurants, for stores within 20 km of a homeless shelter. POST is an indicator taking on a value of 1 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/100, of an establishment to the nearest homeless shelter. In Panel B, distance is variously measured as bins for 0–2, 2–5, 5–10, 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 2 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. *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

	Dependent Variable: log(MONTHLY_STORE_VISITS)		
Panel A. Baseline Regressions of Shelter Distance			
STARBUCKS × POST × DISTANCE	0.330**	0.329**	0.310**
	(2.29)	(2.28)	(2.16)
RESTAURANT × POST × DISTANCE	−0.293***	−0.294***	−0.297***
	(−3.92)	(−3.92)	(−3.95)
STARBUCKS × POST	−0.089***	−0.089***	−0.088***
	(−6.60)	(−6.61)	(−6.39)
RESTAURANT × POST	0.020	0.020	0.019
	(1.13)	(1.13)	(1.04)
POST × DISTANCE	0.084	0.085	0.040
	(0.84)	(0.85)	(0.34)
POST	0.033*		
	(1.99)		
Store-fixed effects	Yes	Yes	Yes
Time-fixed effects	No	Yes	No
City by time-fixed effects	No	No	Yes
No. of obs.	2,701,258	2,701,258	2,701,246
R^2	0.921	0.921	0.922
Panel B. Alternative Versions of Homeless Shelter Distance Regressions			
STARBUCKS × POST × (DIST < 2)	−0.085***		
	(−5.28)		
STARBUCKS × POST × (2 < DIST < 5)	−0.077***		
	(−5.41)		
STARBUCKS × POST × (5 < DIST < 10)	−0.065***		
	(−4.56)		
STARBUCKS × POST × (10 < DIST < 20)	−0.048**		
	(−2.52)		
STARBUCKS × POST × (DIST_Q1)		−0.086***	
		(−5.45)	
STARBUCKS × POST × (DIST_Q2)		−0.086***	
		(−5.81)	
STARBUCKS × POST × (DIST_Q3)		−0.070***	
		(−4.49)	
STARBUCKS × POST × (DIST_Q4)		−0.062***	
		(−4.08)	
STARBUCKS × POST × (DIST_Q5)		−0.051***	
		(−2.92)	
STARBUCKS × POST × log(DISTANCE)			0.012*
			(2.05)
Store-fixed effects	Yes	Yes	Yes
City by time-fixed effects	Yes	Yes	Yes
F-test for [SB × POST × CLOSE] = [SB × POST × FAR]	3.76	3.19	NA
p-Value	0.0662	0.0887	NA
No. of obs.	2,701,246	2,701,246	2,701,246
R^2	0.922	0.922	0.922

(continued on next page)

TABLE 3 (continued)

Starbucks Visits and Distance to Homeless Shelters After Bathroom Policy Change

Panel C. Urban Density and Shelter Distance

	Dependent Variable: log(MONTHLY_STORE_VISITS)			
STARBUCKS × POST	0.027 (0.79)	0.007 (0.23)	−0.007 (−0.23)	−0.028 (−0.91)
STARBUCKS × POST × DISTANCE			0.279* (2.02)	0.261* (1.82)
STARBUCKS × POST × VISITOR_DENSITY	−0.119*** (−2.96)		−0.094** (−2.46)	
STARBUCKS × POST × STORE_DENSITY		−0.098** (−2.65)		−0.071* (−2.00)
Store-fixed effects	Yes	Yes	Yes	Yes
Time-fixed effects	No	No	No	No
City by time-fixed effects	Yes	Yes	Yes	Yes
No. of obs.	2,701,246	2,686,789	2,701,246	2,686,789
R^2	0.922	0.922	0.922	0.922

All of these effects are individually significant at the 1% level, with the exception of the subset farthest from a homeless shelter. 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 3.76, corresponding to a p-value of 0.0662.

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 8.0% relative to other coffee shops after the policy change, whereas stores in the furthest quintile experienced a decline of 5.3%. The difference between these two coefficients has a p-value of 0.0887. Finally, in column 3 we use the same test as in Panel A of Table 3, but replace linear distance with the natural log of distance. The effect is still evident, with a coefficient of 0.012 and a t-statistic of 2.05. 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 a 77% 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, 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. The first is the density of the 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 of Table 3, 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 −0.119 (with a *t*-stat. of −2.96) means that the densest zip code sees an 11.9% reduction in visitors relative to the least dense (the STARBUCKS × POST coefficient, here an insignificant 2.7% increase). The effect of store density in column 2 is slightly smaller, with a coefficient of −0.098 (*t*-stat. of −2.65). 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 0.279 (*t*-stat. of 2.02) and 0.261 (*t*-stat. of 1.82) respectively after controlling for visitor and store density, respectively. This is very similar to the equivalent coefficient in column 3 of Panel A of 0.310, meaning that controlling for urban density makes little difference to the estimated effect of distance from a homeless shelter. The coefficients on density are reduced slightly for visitor density (−0.094), and store density (−0.071), though again the effect is not large. This suggests that the two metrics are capturing largely separate effects.

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.

Homeless shelter distance mattering is a direct prediction of a bathroom channel, but not an obvious prediction of a lingering bad reputation channel. The magnitude of the difference between close-to and far-from-shelter Starbucks locations is 3.7% (8.5%–4.8%). This is roughly half of the baseline 7.0% decline observed in Table 2. This is consistent with a large direct negative effect of the bathroom policy, especially through its effect on the homeless. For there to not be a bathroom channel, there must be some other lingering negative reputation effect that is also correlated with the distance to a homeless shelter, even after controlling for related geographic measures like urban density. This is possible, but not obvious.

C. 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 nontreated 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. Moreover, it cannot be driven by either a taste shock for coffee around the homeless shelter, or a correlated change in preferences near dense urban areas which we consider in Panel C of Table 3. 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 4 distance bins described in the first specification in Panel A of Table 3. 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

FIGURE 2

Trends in Starbucks Versus Other Coffee Shops Before Policy Change, Split by Distance

Figure 2 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 Oct. 2018 levels. This is then averaged by month, store category (Starbucks vs. Coffee Shops), and binned distance from homeless shelter (0–2, 2–5, 5–10, and 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.

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.

To account for the potential pretreatment deviation of foot traffic to Starbucks relative to other coffee shops, we use 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 nontreated observations (non-Starbucks coffee shops) that most closely resembles the treated observation in the pretreatment 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.[17] 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.[18]

We reconsider the analysis performed in the previous table under this alternative framework, with results presented 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. In relative terms, the importance of close-to-shelters stores versus far-from-shelters stores is much larger under synthetic controls. When considering the synthetic control method, the overall reduction is almost exclusively driven by nearby stores, as the point estimate for stores in the farthest bin decreases substantially and becomes statistically insignificant. 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.

D. 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 income and white resident shares and run similar tests to those in Table 2 to find out if Starbucks experienced different demographic changes relative to other coffee shops.

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

[18] 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.

https://doi.org/10.1017/S0022109022001442 Published online by Cambridge University Press

https://doi.org/10.1017/S0022109022001442 Published online by Cambridge University Press

TABLE 4

Robustness for Distance Results Using Synthetic Control Method

Table 4 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 Jan. 2017 and Oct. 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 pretreatment outcome variables for each of the treated observations. For each treated observation, we restrict the sample of candidate nontreated observations to coffee shops within the same 3-digit ZIP code. POST is an indicator taking on a value of 1 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. *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

	Dependent Variable: log(MONTHLY_STORE_VISITS)			
STARBUCKS × POST × LINEAR_DISTANCE	0.462***			
	(4.46)			
STARBUCKS × POST × log(DISTANCE)		0.020***		
		(4.90)		
STARBUCKS × POST × (DIST < 2)			−0.046***	
			(−5.87)	
STARBUCKS × POST × (2 < DIST < 5)			−0.018*	
			(−1.85)	
STARBUCKS × POST × (5 < DIST < 10)			−0.008	
			(−0.68)	
STARBUCKS × POST × (10 < DIST < 20)			0.013	
			(0.98)	
STARBUCKS × POST × (DIST_Q1)				−0.050***
				(−4.88)
STARBUCKS × POST × (DIST_Q2)				−0.033***
				(−3.59)
STARBUCKS × POST × (DIST_Q3)				−0.011
				(−0.92)
STARBUCKS × POST × (DIST_Q4)				−0.010
				(−0.93)
STARBUCKS × POST × (DIST_Q5)				0.012
				(0.95)
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]	NA	NA	18.60	18.02
p-Value	NA	NA	0.0003	0.0004
No. of obs.	2,516,821	2,516,821	2,516,821	2,516,821
R^2	0.939	0.939	0.939	0.939

TABLE 5

Income of Customers After Bathroom Policy Change

Table 5 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 a random sample of noncoffee 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. *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

	Dependent Variable: log(CUSTOMER_ZIP_CODE_INCOME)		
STARBUCKS × POST	−0.003**	−0.003**	−0.004***
	(−2.26)	(−2.13)	(−2.90)
RESTAURANT × POST	−0.006***	−0.006***	−0.006***
	(−3.64)	(−3.72)	(−4.38)
POST	0.008***		
	(2.86)		
Time-fixed effects	No	Yes	No
Store-fixed effects	Yes	Yes	Yes
City by time-fixed effects	No	No	Yes
No. of obs.	2,270,137	2,270,137	2,270,017
R^2	0.855	0.855	0.857

TABLE 6

Race of Customers After Bathroom Policy Change

Table 6 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 a random sample of noncoffee 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. *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

	Dependent Variable: CUSTOMER_ZIP_CODE_PCT_WHITE		
STARBUCKS × POST	0.000 (0.67)	0.001 (1.21)	0.001 (1.21)
RESTAURANT × POST	−0.001* (−1.86)	−0.001* (−1.90)	−0.001** (−2.13)
POST	−0.003*** (−2.88)		
Time-fixed effects	No	Yes	No
Store-fixed effects	Yes	Yes	Yes
City by time-fixed effects	No	No	Yes
No. of obs.	2,279,313	2,279,313	2,279,196
R^2	0.941	0.941	0.942

The income measure relates to the hypothesis that wealthier customers may have different preferences associated with increased crowding in coffee shops, or different preferences over associating with the homeless. The race measure relates to the fact that the original controversy centered in part around allegations that there was racial bias in the enforcement of the previous customers-only bathroom policy.[19] The publicity from the policy change may have either highlighted perceptions that Starbucks was previously acting in a racist manner, or, conversely, may have resulted in a greater appreciation for a policy change aimed at being more racially sensitive. Such effects, however arising, may plausibly have a different impact on black and white customers.

The results for income are presented in Table 5. The dependent variable is the average income of customers. The average income of Starbucks customers decreased relative to the average income of other coffee shops over the treatment period. The coefficients are between −0.003 and −0.004 regardless of the fixed effects used, with *t*-statistics between −2.26 and −2.90. This indicates that the decline in Starbucks attendance documented in previous tables was greater among its wealthier clientele. This is consistent with them being more sensitive to crowding and the new visitors brought in by the bathroom policy.

In Table 6, we do the same tests as above, but replace 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 does not speak to either Starbucks motivation for the original policy, nor the original policy's effect on different racial groups. Rather, it is informative of the attitudes of different races

[19]See, for instance, https://cnn.it/3uOE4tZ. Relatedly, Starbucks closed their stores for an afternoon so employees could undergo training in understanding racial biases.

https://doi.org/10.1017/S0022109022001442 Published online by Cambridge University Press

of customers toward 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 nonwhite Starbucks customers.

E. Time Spent in Store

While the results reported so far inform us on the extensive margin of the treatment (i.e., the number of customer visits to an establishment), they are silent on the intensive margin (i.e., how much money visiting customers spent in an establishment). Because we do not have store-level revenues, we rely on a measure that our cell phone data can provide, namely time spent in the store. As we discussed in the data section, devices periodically ping the current location, with time between pings ranging from a few seconds to several minutes. By using the time between pings and the location of the device during consecutive 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 visit following the policy enactment.

Changes in time spent in the stores are also informative of a direct bathroom channel. Under a lost reputation explanation for the negative effects (both overall, and for close-to-homeless-shelter stores specifically), the simplest explanation is that dislike of Starbucks is deterring previous customers due to bad reputation effects. However, the simplest version of this is that prior customers who now dislike Starbucks will simply avoid the store altogether. It is not clear why they should turn up to Starbucks anyway, but then stay for a shorter period of time. On the other hand, this is a concrete prediction of a bathroom channel, where customers turn up and find that the quality of the amenities is no longer to their liking and decide to leave. As a result, time spent in the store by those customers that *do* still arrive is further evidence of a direct bathroom effect.

TABLE 7

Time Spent in Store After Bathroom Policy Change

Table 7 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 a random sample of noncoffee 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. *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

	Dependent Variable: AVG_MIN_SPENT_IN_STORE		
STARBUCKS × POST	−0.041***	−0.041***	−0.041***
	(−3.84)	(−3.73)	(−5.20)
RESTAURANT × POST	0.035***	0.035***	0.042***
	(6.81)	(6.61)	(7.04)
POST	−0.029		
	(−1.28)		
Time-fixed effects	No	Yes	No
Store-fixed effects	Yes	Yes	Yes
City by time-fixed effects	No	No	Yes
No. of obs.	3,680,074	3,680,074	3,679,978
R^2	0.784	0.786	0.789

https://doi.org/10.1017/S0022109022001442 Published online by Cambridge University Press

The results are presented in Table 7. In the full specification with store and city-by-month fixed effects, customers spend on average 4.1% less time in Starbucks relative to other coffee shops after the policy change, with a *t*-statistic of 5.20. In other specifications, we continue to find similar results, which suggests that not only did fewer customers visited Starbucks stores, but the 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 nonpaying visitors who *do* linger and use tables and bathrooms has an outsized effect on the total number of visitors, who either stop coming and/or spend less time in the store.

Without information on how much customers spent when they visit a given store, we cannot speak directly to the overall revenue impact of the policy enactment. However, given the reduction in both extensive and intensive margins, it seems unlikely that enacting the completely open public bathroom has benefited Starbucks unless the customers increased their purchases 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.

F. Coffee Consumption at Home

In our next set of tests, we investigate the effect of Starbucks' new policy on coffee-related demand satisfied through other channels (e.g., brewed at home). For this purpose, we use Nielsen Homescan Data (available through the Kilts Center at the University of Chicago Booth School of Business) to obtain a measure of coffee purchases. This database provides detailed household-level food purchase information on various attributes such as price, quantity, store location, and purchase date. The households that participate in the data collection process use in-home scanners to record their purchases. We begin by identifying all purchases of products with the term "coffee" in the product category from noncoffee shop retailers.[20] We restrict our sample to households in the yearly panel from 2016 to 2019.

Next, we map the SafeGraph information on device home census block group (used to perform the previous tests for a change in the composition of customers in Table 6) to home zip codes. Using this zip code-level visit data, we construct proxies for the relative popularity of Starbucks compared to other coffee shops using visit data prior to the policy change. With this, we test whether Nielsen panel participants from zip codes who frequent more Starbucks locations saw a relative increase in (noncoffee shop) coffee purchases relative to those zip codes less likely to visit Starbucks, using the following OLS regression:

[20]The Nielsen data reports anonymized store brands IDs, but also the store's industry. As such, we exclude purchases from stores classified as coffee shops, as this would include coffee purchased from Starbucks and other coffee shop competitors. Most purchases with the term "coffee" are from the "Ground and Whole Bean Coffee" category (60.2%).

TABLE 8

Coffee Consumption at Home

Table 8 reports the results of OLS regressions where the dependent variable is the weekly log of weekly household coffee spending from noncoffee shop sources, using data from Nielsen. The independent measures are the interaction of POST (a dummy for periods after the Starbucks bathroom policy change) with measures of Starbucks' popularity at the zip code level. Columns 1–3 measure the share of unique Starbucks stores visited relative to other coffee shops (as levels, percentiles, and terciles, respectively). Columns 4–6 measure visit counts to Starbucks relative to other coffee shops in the preperiod. Household and date-fixed effects are included, and standard errors are double-clustered at the zip code and date level. *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels.

	Dependent Variable: log(HOUSEHOLD_RETAIL_COFFEE_SPENDING)					
SB_STORE_COUNT × POST	0.077** (2.04)					
SB_STORE_COUNT_PCTILE × POST		0.064** (2.05)				
SB_STORE_COUNT_TERCILE_2 × POST			0.020 (0.98)			
SB_STORE_COUNT_TERCILE_3 × POST			0.042* (1.97)			
SB_VISIT_COUNT × POST				0.052** (2.03)		
SB_VISIT_COUNT_PCTILE × POST					0.064** (2.11)	
SB_VISIT_COUNT_TERCILE_2 × POST						0.002 (0.14)
SB_VISIT_COUNT_TERCILE_3 × POST						0.052** (2.14)
Household-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Date-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
No. of obs.	8,096,422	8,096,422	8,096,422	8,096,422	8,096,422	8,096,422
R^2	0.037	0.037	0.037	0.037	0.037	0.037

$$\text{(3)} \quad \text{COFFEE\_PURCHASES}_{ijt} = \beta 1 \times \text{POST} \times f\left(\text{SB\_POP}_{ij}\right) + \text{HOUSEHOLD\_FIXED\_EFFECTS} + \text{TIME\_FIXED\_EFFECTS},$$

where $\text{COFFEE\_PURCHASES}_{ijt}$ is the weekly dollar amount of coffee purchases of household i, from zip code j, in week t. We normalize each household's purchases by their average weekly expenditure in the preperiod. SB_POP_{ij} represents the proxy for Starbucks' popularity, based on the share of unique Starbucks stores visited versus all unique coffee shops on either an equal-weighted or visit-weighted basis. We consider multiple functional forms for the effect of each proxy, including a linear effect of the proxy, linear in the percentile of the proxy, and terciles.

The results reported in Table 8 suggest that after the policy change, households from zip codes that frequent more Starbucks locations in the preperiod saw a significant increase in coffee purchases from noncoffee shops relative to those zip codes where Starbucks was less popular. Each of the [SB_POPULARITY] × POST variables is significant at the 5% level or greater, except the tercile split of SB store counts, which is significant at the 10% level. This is consistent with the idea that previous Starbucks customers were more likely to substitute away from in-store purchases toward more home coffee consumption. In terms of magnitudes, a 1-standard-deviation increase in Starbucks visit counts (0.345) leads to a 1.8% increase in at home coffee consumption after the policy shift (0.052×0.345, from column 4). The magnitudes for a 1-standard-deviation change in the percentile measure from column 5 are similar (1.9% increase).

G. Public Urination Consequences

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 microdata reported by 3 cities (i.e., Austin, Denver, and Pittsburgh) on public urination-related crimes. The idea here is that people urinating 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 or city-by-month fixed effects, to strip out general time-series changes in crime rates.

These results are presented in Panel A of Table 9. 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 1-standard-deviation further from a Starbucks (1.19 km) will have public urination citations in the post-period that are higher by 24.5% ($0.206 \times 1.19 = 0.245$) relative to the closer block group, indicating an economically large effect.

These patterns may simply reflect a general change in criminal activity in these areas, however, and not public urination specifically. To test this, in Panel B of Table 9 we conduct the same full specification tests (columns 2 and 4) for a range of placebo citations in the same 3 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 3 cities reporting nontrivial numbers of geocoded crimes with a sufficiently similar description to be able to be classified.[21] All numbers of citations are again scaled by the block group average, so the coefficients are of comparable magnitude across the different crime types.

[21]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, and so on.

TABLE 9

Crime Related to Public Urination After Bathroom Policy Change

Table 9 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 geocoded crime incident locations), from Jan. 2016 to Dec. 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 is considered – disturbing the peace, simple assault/fighting, marijuana possession, shoplifting, theft of service, threats and harassment, and vandalism. These are chosen based on having nontrivial numbers of classifiable crimes of similar description in all 3 cities. Reported *t*-statistics in parentheses are heteroscedasticity-robust and double-clustered by block group and month. *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A. Public Urination Citations

	Dependent Variable: 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
No. of obs.	12,600	12,600	12,600	12,600
R^2	0.003	0.011	0.003	0.011

Panel B. Public Urination Versus Other Crimes

	Public Urination	Disturbing the Peace	Assaults, Fighting	Marijuana Possession	Shop-lifting	Theft of Service	Threats and Harassment	Vandalism
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
No. of obs.	12,600	11,520	10,656	8,856	7,344	6,336	11,592	10,260
R^2	0.011	0.017	0.012	0.020	0.019	0.022	0.018	0.019
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
No. of obs.	12,600	11,520	10,656	8,856	7,344	6,336	11,592	10,260
R^2	0.011	0.017	0.012	0.020	0.019	0.022	0.018	0.019

Panel B of Table 9 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 the 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

https://doi.org/10.1017/S0022109022001442 Published online by Cambridge University Press

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 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.

H. Financial Market Effects

As an additional confirmation of the economic importance of the detailed store-level analysis, we examine the aggregate effects on Starbucks' financial metrics. We place less emphasis on this analysis, primarily because it is difficult to get a clean counterfactual group, and the outcome variable is aggregated to the public company level, rather than the much finer establishment level. The coffee chains that seem to be the closest counterfactuals are not publicly traded in this period (Peet's, Coffee Bean and Tea Leaf, Caffe Nero, etc.). The closest comparison group of coffee-related public companies is comprised of Dunkin' Donuts, Restaurant Brands International (which owns Tim Hortons), Keurig Doctor Pepper, and J.M. Smucker (which owns Folgers and a number of other coffee brands). We compare Starbucks to an equal-weighted portfolio of these other coffee-related companies, and subtract off the value-weighted market return over the same period. The important effects we document are sufficiently large that the choice of adjustment does not matter here.

Starbucks' stock returns are consistent with the large negative effects we document, but unfortunately not in a manner that enables us to distinguish between different channels. On the day immediately after the incident itself (Apr. 16, 2018), market-adjusted returns of both groups did not indicate a perception of a large problem (−0.48% for Starbucks, but −1.35% for other coffee chains). By the end of the first week, which encompassed a lot of the initial protests, Starbucks was slightly below other coffee shops in cumulative market-adjusted returns (−2.79% vs. −2.38% for other coffee shops). On May 10th, the earliest news story we could find discussing the new bathroom policy, the daily market-adjusted returns were again similar (0.21% for Starbucks, −0.22% for other coffee shops). Starbucks broadly underperformed over the 2 months following the incident, such that by June 19, 2018, Starbucks had a cumulative market-adjusted return of −7.21%, versus −1.96% for other coffee shops.

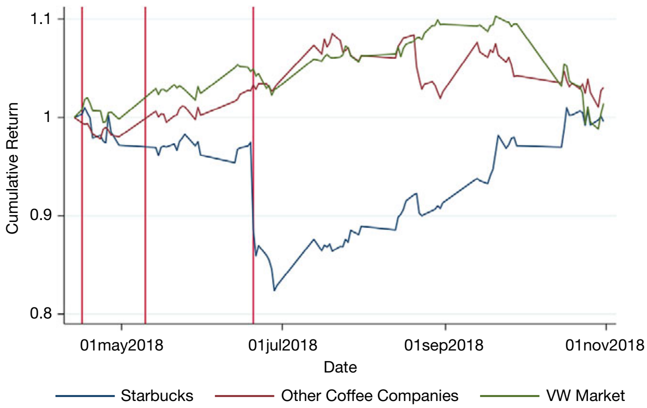
The largest reaction, however, was on June 20, 2018, when markets responded to negative sales and earnings guidance from Starbucks about their recent performance (along with other disclosures in the same press release and conference call, thus muddying the precise inference). Starbucks experienced 2-day market-adjusted returns of −11.12% versus 0.58% for other coffee shops. Of the 35 analysts who revised full-year sales forecasts on June 19th or 20th, the median reduction in full-year sales forecasts was $142 million. This seems to suggest that the market had not anticipated the full effect of the events on future sales until they received guidance from Starbucks. Consequently, it is hard to know from the timing of this response how much of this reaction is to the protests themselves, versus a bathroom channel.

https://doi.org/10.1017/S0022109022001442 Published online by Cambridge University Press

FIGURE 3

Stock Returns for Starbucks, Other Coffee-Related Companies, and the Market

Figure 3 plots the cumulative gross stock returns from Apr. 16, 2018 to Oct. 31, 2018 for 3 groups – Starbucks (in blue), an equal-weighted portfolio of 4 coffee-related companies (Dunkin' Donuts, Restaurant Brands International (which owns Tim Hortons), Keurig Doctor Pepper, and J.M. Smucker (a consumer goods company)), and the value-weighted market portfolio. The 3 red lines correspond to (respectively) the first trading day after the initial incident (Apr. 16, 2018), the first news report of the new bathroom policy (May 10, 2018), and the trading day when Starbucks issued a press release, 8-K and conference call with negative earnings guidance (June 20, 2018).



A large amount of this negative stock return for Starbucks eventually recovered, but even by the end of Oct. 2018 when our sample ends, Starbucks had adjusted market returns of −1.85%, versus 1.64% for other coffee shops. We show the cumulative returns from Apr. 16, 2018 until Oct. 31, 2018 (the end of our store sample) in Figure 3. The 3 vertical lines correspond to the date of the initial incident, the first news of the change in bathroom policy, and the date of the major negative earnings revisions for the second quarter.

IV. Conclusion

While the determination of the appropriate level of CSR engagement is generally discussed using a firm profitability litmus test, the viability of incremental CSR activity also relies on a manager's ability to aggregate preferences across investor and noninvestor stakeholders and the clear generation of a net social benefit. The Starbucks experiment provides an opportunity to evaluate each tenet underlying stakeholder capitalism. Our evidence suggests that the Starbucks policy decreased both the foot traffic and time spent at Starbucks, especially in stores that are closer to homeless shelters.

Beyond the impact on firm profitability, our results highlight the difficulty in pursuing ESG activities that provide an amenity to noncustomers. The inability of this group to provide market signals complicates an inference of the marginal benefit they receive, hampering a manager's ability to aggregate preferences across heterogeneous stakeholders.

Moreover, while certain amenities 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 amenity to nonpatrons, who may also actively deter others' purchases. In such instances, the leftward shift in the demand curve for products stemming from the provision of the public amenity suggests a reduction in consumer surplus, potentially offsetting the social benefit associated with the public amenity.

Finally, the lessons we learn from Starbucks' experiment can be generalized to other domains. It is not uncommon to see governments seeking help from private companies to provide public goods such as national parks, national defense, public broadcasting, clean air, or space exploration. This is perhaps not that surprising because the profits of private corporations often dwarf the economy of many countries across the globe, and also many U.S. states. As corporations fill the void of public good provision, it is perhaps inevitable to face conflicts with their main mandate, that is, profit maximization for their shareholders. Our results suggest a particular cost of certain types of CSR strategies that involve making prosocial behavior a core part of their business. In this case, the prosocial behavior comes at the cost of actually impeding the core business of the company. Our results suggest that companies may be better off focusing on donating money to worthwhile causes, and effectively using a division of labor, whereby Starbucks specializes in making and selling coffee, and engages in CSR by supporting organizations who specialize in social policies. Our results show that trying to incorporate the two within a single company may result in outsized negative externalities for the underlying business that makes CSR possible in the first place.

Appendix

Starbucks reports quarterly sales growth and revenue numbers for the first three fiscal quarters, and full-year numbers for the fourth fiscal quarter. Fourth-quarter revenue is computed as the difference between full-year revenue and the sum of revenue for the previous three quarters. For fourth-quarter sales growth, 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:

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

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

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

Thus we have

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

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

References

Abadie, A.; A. Diamond; and J. Hainmueller. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association*, 105 (2012), 493–505.

Abadie, A., and J. Gardeazabal. "The Economic Costs of Conflict: A Case Study of the Basque Country, *American Economic Review*, 93 (2003), 113–132.

Athey, S.; B. Ferguson; M. Gentzkow; and T. Schmidt. "Experienced Segregation." Working Paper, Stanford University (2019).

Becker-Olsen, K. L.; B. A. Cudmore; and R. P. Hill. "The Impact of Perceived Corporate Social Responsibility on Consumer Behavior." *Journal of Business Research*, 59 (2006), 46–53.

Bénabou, R., and J. Tirole. "Individual and Corporate Social Responsibility." *Economica*, 77 (2010), 1–19.

Bhattacharya, C. B.; S. Sen; and D. Korschun. "Using Corporate Social Responsibility to Win the War for Talent." *MIT Sloan Management Review*, 49 (2008), 37–44.

Bode, C.; J. Singh; and M. Rogan. "Corporate Social Initiatives and Employee Retention." *Organization Science*, 26 (2015), 1702–1720.

Brekke, K. A., and K. Nyborg. "Attracting Responsible Employees: Green Production as Labor Market Screening." *Resource and Energy Economics*, 30 (2008), 509–526.

Chen, M. K., and R. Rohla, "The Effect of Partisanship and Political Advertising on Close Family Ties." *Science*, 360 (2018), 1020–1024.

Cheng, H.; H. Hong; and K. Shue. "Do Managers Do Good with Other People's Money?" NBER Research Paper No. w19432 (2013).

Cheng, B.; I. Ioannou; and G. Serafeim. "Corporate Social Responsibility and Access to Finance." *Strategic Management Journal*, 35 (2014), 1–23.

Dai, R.; H. Liang; and L. Ng. "Socially Responsible Corporate Customers." *Journal of Financial Economics*, 142 (2021), 598–626.

Dhaliwal, D.; O. Z. Li; A. H. Tsang; and Y. G. Yang. "Voluntary Non-Financial Disclosure and the Cost of Equity Capital: The Case of Corporate Social Responsibility Reporting." *Accounting Review*, 86 (2011), 59–100.

Eccles, R. G.; I. Ioannou; and G. Serafeim. "The Impact of Corporate Sustainability on Organizational Processes and Performance." *Management Science*, 60 (2014), 2835–2857.

Edmans, A. "Does the Stock Market Fully Value Intangibles? Employee Satisfaction and Equity Prices." *Journal of Financial Economics*, 101 (2011), 621–640.

El Ghoul, S.; O. Guedhami; C. C. Y. Kwok; and D. R. Mishra. "Does Corporate Social Responsibility Affect the Cost of Capital?" *Journal of Banking and Finance*, 35 (2011), 2388–2406.

Ferrell, A.; H. Liang; and L. Renneboog. "Socially Responsible Firms." *Journal of Financial Economics*, 122 (2016), 585–606.

Flammer, C. "Does Corporate Social Responsibility Lead to Superior Financial Performance? A Regression Discontinuity Approach." *Management Science*, 61 (2015), 2549–2568.

Friedman, M. "The Social Responsibility of Business is to Increase its Profits." *New York Times Magazine*, Sept. 13, https://www.nytimes.com/1970/09/13/archives/article-15-no-title.html (1970).

Gormley, T. A., and D. A. Matsa. "Common Errors: How to (and not to) Control for Unobserved Heterogeneity." *Review of Financial Studies*, 27 (2014), 617–661.

Greening, D. W., and D. B. Turban. "Corporate Social Performance as a Competitive Advantage in Attracting a Quality Workforce." *Business & Society*, 39 (2000), 254–280.

Gregg, P; P. A. Grout; A. Ratcliffe; S. Smith; and F. Windmeijer. "How Important is Pro-Social Behaviour in the Delivery of Public Services?" *Journal of Public Economics*, 95 (2011), 758–766.

Hart, O., and L. Zingales. "Companies Should Maximize Shareholder Welfare Not Market Value." *Journal of Law, Finance, and Accounting*, 2 (2017), 247–275.

Hwang, S. W.; E. Gogosis; C. Chambers; J. R. Dunn; J. S. Hoch; and T. Aubry. "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 (2011), 1076–1090.

Imbens, G. W., and J. M. Wooldridge. "Recent Developments in the Econometrics of Program Evaluation." *Journal of Economic Literature*, 47 (2009), 5–86.

Institute of Medicine (US) Committee on Health Care for Homeless People. *Homelessness, Health, and Human Needs*. Washington, DC: National Academies Press (1988). Available from: https://www.ncbi.nlm.nih.gov/books/NBK218232/.

Jones, D. A.; C. R. Willness; and S. Madey. "Why are Job Seekers Attracted by Corporate Social Performance? Experimental and Field Tests of Three Signal-Based Mechanisms." *Academy of Management Journal*, 57 (2014), 383–404.

Karpoff, J. M. "On a Stakeholder Model of Corporate Governance." *Financial Management*, 50 (2021), 321–343.

Karpoff, J. M.; D. S. Lee; and G. S. Martin. "The Cost to Firms of Cooking the Books." *Journal of Financial and Quantitative Analysis*, 43 (2008), 581–612.

Kim, Y.; M. Park; and B. Wier. "Is Earnings Quality Associated with Corporate Social Responsibility?" *Accounting Review*, 87 (2012), 761–796.

Koh, P.-S.; C. Qian; and H. Wang. "Firm Litigation Risk and the Insurance Value of Corporate Social Responsibility." *Strategic Management Journal*, 35 (2014), 1464–1482.

Krüger, P. "Corporate Goodness and Shareholder Wealth." *Journal of Financial Economics*, 115 (2015), 304–329.

Kushel, M. B.; S. Perry; D. Bangsberg; R. Clark; and A. R. Moss. "Emergency Department Use Among the Homeless and Marginally Housed: Results From a Community-Based Study." *American Journal of Public Health*, 92 (2002), 778–784.

Liang, H., and L. Renneboog. "On the Foundations of Corporate Social Responsibility." *Journal of Finance*, 72 (2017), 853–910.

Liang, H., and L. Renneboog. "Corporate Social Responsibility and Sustainable Finance: A Review of the Literature." Finance Working Paper No. 701/2020, European Corporate Governance Institute (2020).

Lins, K. V.; H. Servaes; and A. Tamayo. "Social Capital, Trust, and Firm Performance: The Value of Corporate Social Responsibility During the Financial Crisis." *Journal of Finance*, 72 (2017), 1785–1824.

Long, E.; K. Chen; and R. Ryne. "Political Storms: Emergent Partisan Skepticism of Hurricane Risks." Working Paper, available at ssrn.com/abstract=3339723 (2019).

Luo, X., and C. B. Bhattacharya. "Corporate Social Responsibility, Customer Satisfaction, and Market Value." *Journal of Marketing*, 70 (2006), 1–18.

Luo, X., and S. Du. "Exploring the Relationship Between Corporate Social Responsibility and Firm Innovation." *Marketing Letters*, 26 (2015), 703–714.

Mackey, A.; T. B. Mackey; and J. B. Barney. "Corporate Social Responsibility and Firm Performance: Investor Preferences and Corporate Strategies." *Academy of Management Review*, 32 (2007), 817–835.

Masulis, R. W., and S. W. Reza. "Agency Problems of Corporate Philanthropy." *Review of Financial Studies*, 28 (2015), 592–636.

Matten, D.; A. Crane; and W. Chapple. "Behind the Mask: Revealing the True Face of Corporate Citizenship." *Journal of Business Ethics*, 45 (2003), 109–120.

Ortiz-de-Mandojana, N., and P. Bansal. "The Long-Term Benefits of Organizational Resilience Through Sustainable Business Practices." *Strategic Management Journal*, 37 (2016), 1615–1631.

Pástor, L.; R. F. Stambaugh; and L. A. Taylor. "Sustainable Investing in Equilibrium." *Journal of Financial Economics*, 142 (2021), 550–571.

Rieke, K.; A. Smolsky; E. Bock; L. P. Erkes; E. Porterfield; and S. Watanabe-Galloway. "Mental and Nonmental Health Hospital Admissions Among Chronically Homeless Adults Before and After Supportive Housing Placement." *Social Work in Public Health*, 30 (2015), 496–503.

Servaes, H., and A. Tamayo. "The Impact of Corporate Social Responsibility on Firm Value: The Role of Customer Awareness." *Management Science*, 59 (2013), 1045–1061.

Sun, W., and K. Cui. "Linking Corporate Social Responsibility to Firm Default Risk." *European Management Journal*, 32 (2014), 275–287.

Vishwanathan, P.; H. (J.) van Oosterhout; P. P. M. A. R. Heugens; P. Duranc; and M. van Essen. "Strategic CSR: A Concept Building Meta-Analysis." *Journal of Management Studies*, 57 (2020), 314–350.