## Title
Social Learning in the COVID-19 Pandemic: Community Establishments' Closure Decisions Follow Those of Nearby Chain Establishments



## Journal
Management Science, 67(7)

## ISSN
0025-1909

## Authors
de Vaan, Mathijs
Mumtaz, Saqib
Nagaraj, Abhishek
et al.

## Publication Date
2021-07-01

## DOI
10.1287/mnsc.2021.4033

Peer reviewed

# Social Learning in the COVID-19 Pandemic: Community Establishments' Closure Decisions Follow Those of Nearby Chain Establishments

Mathijs de Vaan, Saqib Mumtaz, Abhishek Nagaraj, and Sameer B. Srivastava

Haas School of Business, University of California, Berkeley *

**Abstract**

As conveners that bring various stakeholders into the same physical space, firms can powerfully influence the course of pandemics such as COVID-19. Even when operating under government orders and health guidelines, firms have considerable discretion to keep their establishments open or closed during a pandemic. We examine the role of social learning in the exercise of this discretion at the establishment level. In particular, we evaluate how the closure decisions of chain establishments, which are associated with national brands, affect those of proximate, same-industry community establishments, which are independently owned or managed. We conduct these analyses using cellphone location tracking data on daily visits to 230,403 U.S.-based community establishments that are co-located with chain establishments affiliated with 319 national brands. We disentangle the effects of social learning from confounding factors by using an instrumental variables strategy that relies on local variation in community establishments' exposure to closure decisions made by brands at the national level. Our results suggest that closing decisions of community establishments are affected by the decisions made by chain establishments: a community establishment is 3.5% more likely to be open on a given day if the proportion of nearby open chain establishments increases by one standard deviation.

*Corresponding Author: Mathijs de Vaan (mdevaan@haas.berkeley.edu). Other authors: saqib@berkeley.edu, nagaraj@berkeley.edu, and sameersriv@berkeley.edu. We thank Adit Jain for excellent research assistance. We thank three anonymous reviewers, the associate editor and department editor of Management Science for their valuable feedback and suggestions. We also thank Paul Gouvard and András Tilcsik for feedback on prior drafts of the paper. The usual disclaimer applies.

## Introduction

The COVID-19 pandemic has highlighted the role of firms as conveners: despite the proliferation of online communication, many business models still rely on bringing employees, customers, and other stakeholders into the same physical space. Research on infectious disease in general and COVID-19 in particular has shown that physical co-location can fuel disease transmission (Anderson et al., 2020). As a result, local and state governments have implemented a wide range of policies and directives to influence whether and how firms remain open or closed during different phases of the COVID-19 pandemic (Benzell et al., 2020; Goolsbee et al., 2020).

For some firms, government closure directives have left their managers or owners with no choice but to shutter their doors; however, many other firms have been able to exercise discretion in their closure decisions. This discretion arises in part because many counties and some states never explicitly ordered firms to shut down, effectively passing responsibility for responding to the COVID-19 pandemic to business owners. Moreover, even when directives are in place, many local and state governments lack adequate resources to monitor and ensure full compliance of all establishments in their jurisdiction. Finally, directives issued by local and state governments are often ambiguous, leaving business owners with latitude to interpret the guidance as they see fit. Given that firms have considerable leeway in making closure decisions and that these choices can determine the shape of a pandemic's trajectory (Akbarpour et al., 2020; Chang et al., 2020), we examine the social dynamics that underpin these decisions.

In particular, we consider the role of interorganizational social learning in firms' closure decisions during the COVID-19 pandemic. In similar fashion to the introduction of disruptive and ambiguous regulations (Kelly, 2003; Baker et al., 2016), the COVID-19 pandemic and the health directives it has spawned have created a context in which firms have been forced to make closure decisions under high levels of uncertainty—for example, about how their customers will react, how their decision will affect their financial position, whether their employees will resist, and so on. During such times of extreme uncertainty, firms may look to one another for guidance on what behaviors to adopt (DiMaggio and Powell, 1983; Coleman, 1990; Rogers, 2010). One reason for firms to emulate the behavior of other firms is that they believe that these firms have superior information about the appropriate response to uncertainty (Pfeffer et al., 1976; Tushman and Romanelli, 1983). But even if the information held by other firms is not obviously superior, the mere act of adopting the behavior of others can reduce perceived uncertainty (Turner, 2010; Balla-Elliott et al., 2020).

We use the term social learning to describe the process by which the behavior of one organization prompts other organizations to adopt the same behavior (Young, 2009; DiMaggio and Garip, 2012). Such adoption may be the result of direct social influence of the manager of one firm on the manager of another firm. Yet firms can also learn from the actions of other firms in indirect ways—for example, a change in customer behavior in response to the closings of other establishments may represent a signal to the focal firm that the appropriate decision is to shutter its doors too. Given the nature of the data available to us, we are not able to pin down the specific channels, such as interorganizational managerial influence, through which this social learning occurs. We do, however, develop an empirical strategy to demonstrate that closure decisions are not made in isolation: when given the discretion, closure decisions of firms tend to rely on the closure decisions of geographically proximate competitors.

We consider two classes of establishments: (1) *chain establishments*, which are either owned by or affiliated with national brands and subject to more centrally defined organizational policies and practices; and (2) *community establishments*, which are independently owned and managed. We examine the extent to which community establishments' closure decisions follow those of nearby chain establishments in the same industry. While prior work on interorganizational learning has mostly focused on the transmission of information through formal and mostly cooperative network connections, we examine how firms change their behavior in response to decisions made by competitors. Although we acknowledge that community establishments can also influence chain establishments' closure decisions, our analytical focus is on the effect of chain establishments on community establishments because community establishments have considerably more discretion in their closure decisions and our empirical strategy (described in greater detail below) can only identify the latter effect.

We conducted this investigation using anonymized cellphone location tracking data on daily visits to 230,403 U.S.-based community establishments that are co-located with chain establishments affiliated with 319 national brands. Rigorously testing our proposition about social learning poses a key identification challenge: firms that are physically proximate do not only respond to each other's behavior but may also be susceptible to shared contextual forces that affect all firms in the vicinity. For example, a community and chain establishment in the same neighborhood might close not because one firm learns from the closure of another but simply because both firms are faced with the same county directives and the same local COVID-19 infection rate. To begin to address these empirical challenges, we estimate the impact of chain closures on community establishments while controlling for fine-grained fixed effects at the level of the county-date, industry-date and zip code. These fixed effects account for many potentially confounding factors such as daily changes in government policy, local infection rates, changes in local demand, and time trends in closure across different industries.

Although these fixed effects address many threats to causal identification, they fall short of the ideal experiment that would examine the impact of randomly-assigned chain establishment closures on proximate, same-industry community establishments. To approximate this ideal design, we leverage the fact that firms with national scale (such as chains) often make centralized closure decisions that apply to all or a majority of their establishments without placing much weight on any particular local conditions. We can therefore evaluate our research question by comparing the responses of otherwise similar establishments that differ only in their local exposure to chain establishments with different national policies. Building on this intuition and the literature on "shift-share" instruments (Derenoncourt, 2019; Goldsmith-Pinkham et al., 2018), we instrument for the closure decisions of a community establishment's rival chain establishments in the same industry and zip by using the average number of closures of other establishments of the same chains in all states except for the community establishment's focal state. Our results suggest that a community establishment is 3.5% more likely to be open on any given day if the proportion of nearby open chain establishments increases by one standard deviation. Taken at face value, this effect is modest as a chain establishment closure affects a single community establishment's closure decision on a given day by only a small amount. That said, when aggregated over a large number of community establishments over several weeks, the cumulative effect we document becomes meaningful and consequential. Finally, while our data do not allow us to isolate the precise mechanism by which chain establishments affect the behavior of community establishments, our results provide robust evidence consistent with firms paying attention to competitors and imitating their behavior. In the following sections, we introduce our data, discuss our empirical strategy and present our findings.

## Data

We conduct our analyses using a data set released by the SafeGraph COVID-19 Data Consortium.[1]. SafeGraph aggregates anonymized location data from numerous applications on about 45 million mobile devices to provide foot traffic patterns at physical places.[2] We focus on service-oriented points-of-interest (POI) such as retail shops, restaurants, movie theaters, and fitness centers. For each POI, in addition to its daily foot traffic, we observe its geographic location, its North American Industry Classification System (NAICS) industry code, and the branded chain it belongs to (if any). Appendix A1 describes further details about the data.

### Sample Construction

Our primary panel spans the period from March 1, 2020 to April 15, 2020—that is, just before the COVID-19 pandemic led state and local governments to begin issuing shelter-in-place directives and for six weeks thereafter. Given that our focus is on establishments that had discretion in their closure decisions, we eliminate ones that were likely deemed "essential," which we define as NAICS categories in which more than 70% of establishments were open during our observation period. This method identifies establishments in industries such as gas stations and grocery stores which were widely considered essential and provides a principled basis to exclude them from our analyses.

The SafeGraph data indicate whether a given establishment is part of a branded chain as well as the name of the brand, if applicable. We label establishments associated with a national brand as chain establishments and those with no such association as community establishments. We limit our focus to national brands that *contained more than 50 establishments and operated in at least 25 states*. This strategy minimizes the risk that closure decisions by chain establishments were driven by local COVID-19 conditions, a critical condition for our instrumental variables strategy. In total, 319 brands (represented by 198,176 unique establishments) qualify as a national brand. Our baseline sample is composed of community establishments that had at least one chain establishment in their 3-digit NAICS industry code and zip code. In total, our sample includes 10,368,135 establishment-

[1] https://www.safegraph.com/covid-19-data-consortium.

[2] Information from establishment-census block group observations with less than five devices are excluded for privacy reasons.

day observations for 230,403 establishments over 45 days, distributed across approximately 12,000 zip codes across the U.S. Appendix A2 contains validation checks for the sample construction.

**Variables**

Table 1 summarizes our key variables. Our dependent variable *Open* is an indicator for whether an establishment was open or closed on a given day. Since the SafeGraph data do not include such an indicator, we developed an algorithm for this purpose. This algorithm relies on past traffic in February 2020 (before people began sheltering in place), the rate of change of visitors on a daily basis and makes adjustments for very small and large places whose closures can be harder to track. Appendix A3 describes this algorithm and includes various validation checks, and reports the robustness of our results to alternative definitions of the *Open* variable.

[TABLE 1 ABOUT HERE]

Our key independent variable, *Prop. Chain Est. Open*, measures the proportion of chain establishments that are open in the same zip code and industry as the focal community establishment. If a community establishment has multiple rival chain establishments, we take the weighted average of these establishments' closing status. To account for the fact that there is likely to be a lag in social learning, we consider the previous day's proportion for the focal day's closure decision. On average, about 59% of an establishment's chain counterparts in the same industry and zip code were open, although there is significant dispersion around the mean, and also reflects data from the first two weeks of March when shelter-in-place orders were yet to be issued across the U.S.

Our instrumental variable strategy depends on a reliable instrument for *Prop. Chain Est. Open*, which we label *National Chain Opening Exposure*. For a given brand-date-state combination, we computed the proportion of stores that were closed across the country, while excluding establishments in the focal state. Then we calculated the weighted average of this statistic for every community establishment based on the brands in the zip code and industry of the community establishment. This metric provides a measure of the extent to which rival chains of the focal community establishment enacted company wide corporate directives to shutter their doors or remain open. We explain this instrument in further detail after we describe other data and summary statistics.

Our models include various control variables. We approximate establishment size by calculating the average number of visitors in February, *Avg. February Traffic*. The median establishment has about 13 visitors on a given day in February. Note that this number is an order of magnitude lower than the true number of visitors because SafeGraph's coverage is limited to only about one sixth the number of devices of the U.S. population. *Prop. Devices At Home* measures the number of devices that, according to Safegraph, do not move out of their home location on a given date. This is a useful proxy for the extent to which individuals in a given area decided to limit their mobility and helps control for demand factors that might affect closure decisions (Boxell et al., 2020; Chiou and Tucker, 2020). Finally, we define an indicator variable, *Shelter in Place*, that accounts for whether "shelter-in-place" orders were in effect at the county level for a given zip code on a given day (based on data from the National Association of Counties).[3]

## Empirical Strategy and Results

### Motivating Example

To provide greater intuition for our empirical strategy, we begin with a motivating example. Figure 1 shows the closing status of chain and community establishments in the fitness center industry in two neighboring zip codes in Collin County, Texas, on March 25, 2020. As indicated by the star icons, the zip code on the left, 75150, had a closed chain establishment (Orange Theory), while the zip code on the right, 75409, had a chain establishment (Anytime Fitness) that was still open. These closing behaviors were consistent with the broad closure pattern of other establishments in these two chains: as of March 25, 78% of Orange Theory establishments had closed, while 66% of Anytime Fitness establishments had closed according to our data. As depicted in the figure and consistent with our theory and empirical strategy, all six community establishments in the vicinity of the closed Orange Theory (Brand A) establishment were closed, while three out of the five community establishments near the open Anytime Fitness were also open. Moreover, any corporate directives issued by Orange Theory or Anytime Fitness at this time were unlikely to be influenced by the specific local conditions in Collin County, which is the main assumption for our exclusion restriction. We turn next to examining whether the pattern observed in this example generalizes across industries and locations.

[3]See https://tinyurl.com/y6sdlgfd.

[FIGURE 1 ABOUT HERE]

**Identifying a Social Learning Effect**

Prior literature on statistically identifying social learning effects demonstrates that mere similarity in behavior between connected actors does not necessarily imply a causal transmission process (Azoulay et al., 2017; Shalizi and Thomas, 2011). In the example above, we proposed that community establishments in zip code 75150 might be more likely to be closed than those in zip code 75409 because of differential decisions made by their nearby rival chain establishments. Yet other factors such as differences in local neighborhood guidelines, local COVID-19 cases, the demographic composition of residents, and local media coverage could produce a similar pattern.

More formally, consider a model in which a community establishment $i$, in industry $n$, and zip code $z$ makes a decision to stay open at time, $t$, represented by the indicator $Open_{inzt}$ (where 0 indicates closed and 1 indicates open). We are interested in the relationship between this variable and $Prop.\ Chain\ Est.\ Open_{nz(t-1)}$, the proportion of chain establishments in industry, $n$, and zip code, $z$, that are open at time, $(t-1)$. In our example above, this variable equals 0 in zip code 75150 and 1 in zip code 75409.

Inspecting the simple correlation between $Open_{inzt}$ and $Prop.\ Chain\ Est.\ Open_{nz(t-1)}$ would be misleading because of the various confounds noted before. To partially address this issue, in our most stringent specification, we include non-parametric time trends in the form of county-by-date fixed effects $\eta_{zt}$. The county-by-date fixed effects account for a range of confounding factors including daily changes in infections, any county-level social distancing guidelines as well as awareness of the importance of social distancing across time. Further, since national patterns of response to the pandemic likely varied by industry and over time, our most stringent specification also includes industry-date fixed effects $\gamma_{nt}$. Finally, to control for time-invariant differences across zip codes within a county—for example, differences in local income, race, and political orientation (Boxell et al., 2020)—we include zip code fixed effects $\theta_z$.

Beyond the sources of variation accounted for by our fixed effects, we anticipate that formal shelter-in-place orders at the local level also affect local establishments' closure decisions. In models without temporal fixed effects, we therefore include an indicator variable, $Shelter\ In\ Place_{zt}$, which is set to 1 if there is a shelter in place policy at time, $t$, in the county to which a zip code, $z$, belongs. In models with time fixed effects, this variable is subsumed by the fixed effect. Similarly, given that establishment size might be associated with differing incentives to remain open or closed, we include a proxy for size: $Avg.\ February\ Foot\ Traffic_i$. Finally, to account for local, time-varying differences in customer mobility, we include the control, $Prop.\ Devices\ At\ Home_{zt}$, which indicates the proportion of customers in zip, $z$, who are following guidelines to limit their mobility by not leaving their residence on date, $t$. Thus, our most stringent OLS specification can be summarized as follows:

$$\begin{aligned} Open_{inzt} = \beta_0 + \beta_1 * Prop.\ Chain\ Est.\ Open_{nz(t-1)} + Avg.\ Feb\ Foot\ Traffic_i + \\ Prop.\ Devices\ at\ Home_{zt} + \gamma_{nt} + \eta_{zt} + \theta_z + \epsilon \end{aligned} \tag{1}$$

**Instrumental Variables Specification**

Despite the inclusion of control variables and various fixed effects, it is still possible that our OLS estimates of the coefficient of interest, $\beta_1$, are biased since there are likely to be other omitted variables that we cannot explicitly control for. For example, health-conscious customers in a particular area might mobilize to urge both chain and community establishments in a particular industry to close. To account for such possibilities, we developed an instrumental variables (IV) identification strategy. Any potential instrument, $Z$, for our endogenous variable, $Prop.\ Chain\ Est.\ Open_{nz(t-1)}$, must satisfy two conditions. First, it must be sufficiently correlated with $Prop.\ Chain\ Est.\ Open_{nz(t-1)}$— i.e., it should significantly affect the probability that a chain establishment temporarily closes. Second, the IV must satisfy the exclusion restriction: it must affect a community establishment's decision to close exclusively via its influence on closing the local branch of the national chain. Formally, the instrument requires that $cov(Open_{inzt}, Z) \neq 0$ and $cov(Z, \epsilon) = 0$.

Our instrument for $Prop.\ Chain\ Est.\ Open_{nz(t-1)}$ is $National\ Chain\ Opening\ Exposure_{nz(t-1)}$, which we denote by $Z_{nz(t-1)}$ for brevity. To understand how $Z_{nz(t-1)}$ is calculated, consider the example in Figure 1. For each industry-zip pair, we calculate the proportion of chain establishments outside of the state of Texas that are open on date, $(t-1)$. Orange Theory (left panel) has 825 establishments outside Texas, of which 181 are open. Meanwhile, Anytime Fitness (right panel) has 1565 establishments outside Texas, of which 526 are open. Since only one chain brand is active in both cases, the instrument, $Z_{nz(t-1)}$, is simply calculated as the proportion of establishments open outside Texas. In Figure 1, this statistic is 0.22 for the zip code 75150 and 0.34 for zip code 75409. In the case of multiple chains, the instrument, $Z_{nz(t-1)}$, is calculated as the weighted average of this statistic across different chains.

While we cannot fully rule out violations of the exclusion restriction, we can evaluate whether corporate closure decisions were orthogonal to observable local market conditions. Doing so allows us to determine whether the common exposure of chain and community establishments to local business or disease conditions interfere with the instrument and act as a confounder in our models. If that is the case, it is possible that we are detecting a "canary in the coal mine" effect, where chain establishments are responding more rapidly to changing local business and disease conditions, perhaps because they have better resources and are more sensitive to legal risk. Note that in this scenario, chain and community establishments exhibit similar closing behavior that is not driven by social learning but by common exposure to local business or disease conditions. We design a series of analyses aimed at determining whether the most plausible common exposure effects drive our results. We include these results in appendix B. Overall, the results of these robustness checks further corroborate our main finding.

**Summary Statistics**

Table 1 provides descriptive statistics for our main variables of interest. The table shows that the unconditional probability of a community establishment being open in our sample period (i.e., March $2^{nd}$ through April $15^{th}$) is 0.431. Note that this mean includes establishment days before shut downs began around March 15. The table also shows that our main independent variable: *Prop. Branch Est. Open*, ranges between 0 and 1, with a mean of 0.59. The mean of our instrument, *National Chain Opening Exposure*, is slightly higher at 0.61. Finally, the descriptive statistics for our three control variables, *Prop. Devices At Home* and *Shelter In Place* are included.

**Empirical Estimates**

In Panel A of Table 2, we report results of OLS regressions. As stated before, we lag the key independent variable by one day (results are robust to longer lags): the probability of a community establishment being open at day $t$ is a function of *Prop. Branch Est. Open* at day $t-1$. And since the variation we observe might be clustered within spatial units, we adjust the standard errors for clustering at the county-level.

Model 1 presents the association between the probability of a community establishment being open and the proportion of chain establishments that are open in the presence of key controls (but without fixed effects). Model 2 adds industry and zip code fixed effects, and Model 3 includes almost 12,000 zip fixed effects as well. Model 4, our most stringent specification, includes zip fixed effects, but now also includes industry and county-level time trend (NAICS-by-date and county-by-date) fixed-effects rather than a common time trend. The inclusion of fixed effects reduces the coefficient estimate as compared to Model 1 that includes only the controls, demonstrating their importance in this setting. The range of effect sizes reported in Models 2-4 suggests that a one standard deviation increase in the proportion of chain establishments (0.356) that are open leads to between a 1.3%-6.4% increase in the probability of a community establishment being open on a given day. To put this number into perspective it is important to note that the effect is estimated at the establishment-day level: even though the effect is modest, it has the potential to accumulate across establishments and over time.[4]

[TABLE 2 ABOUT HERE]

Next we move to our IV estimates, which are reported in Panel B. The first stage is positive and significant, suggesting that our instrument is predictive of a chain establishment temporarily closing. Also, the first-stage F-statistic is about 1,250, indicating that our instrument is sufficiently powered (Lee et al., 2020). The IV estimates are broadly consistent with our OLS estimates in the sense that they show a positive and significant effect of chain establishments' closure decisions on those of community establishments. If anything, the IV estimates suggest a slightly greater effect size than revealed by the OLS specifications. The IV estimate in Model 6 suggests that a community establishment is 3.5% more likely to be open if the proportion of open chain establishments increases by one standard deviation.

[4] Another way to interpret the effect size is to consider the coefficient of the *Shelter in Place* variable. A change to a Shelter in Place order is associated with a 3.8% reduction in the likelihood of a community establishment being open on a given day.

## Discussion

The goal of this article has been to identify the role of social learning in the closure decisions of community establishments during the COVID-19 pandemic. We argued that these establishments will learn about whether to remain open or to instead close from their proximate chain establishment competitors. Using granular, time-varying data on individuals' visits to a national population of establishments in the U.S. and a novel instrumental variables strategy, we found support for this proposition.

Findings from this study contribute to our understanding of interorganizational social learning. Prior work on interorganizational social learning has focused on the transmission of information through formal and mostly cooperative network connections–for example, in the form of joint ventures, alliances, and interlocking boards (Davis, 1991; Powell et al., 1996). In contrast, we highlight a setting in which firms learn from and adopt the behaviors of geographically proximate competitors. In other words, economically relevant information can be transmitted between organizations in the absence of a formal network connection and even when the two organizations are fierce competitors.

This paper also makes noteworthy methodological contributions. Our paper is one of the first to apply a shift-share instrument to the study of interorganizational social learning. Many previously used estimation strategies cannot account for unobserved, time-varying contextual factors that might affect organizational behavior and that might otherwise masquerade as social learning. The instrument we introduce can be readily extended to study other ways in which centralized decisions of chain establishments might shape the behavior of proximate community establishments. Second, we develop and validate a methodology for determining, based on location tracking data, when a given establishment is open or closed. Given the ubiquity of location tracking data, an approach such as ours can be readily extended to studying other forms of interorganizational social learning—for example, how choices about hours of operation might cascade from one organization to others.

Our findings also have implications for the design of government policies to influence firm behavior in the management of pandemics such as COVID-19. Perhaps most importantly, this paper shows that when government directives and health guidelines are ambiguous, firms will look for other information to guide their decision making. Obviously, such ambiguity may have been intentional if local governments believe that firms are well positioned to make these important decisions. But if one assumes that this is not the case, policy makers and local governments should consider the consequences of a lack of clarity and precision in their directives.

These contributions notwithstanding, the study also has certain limitations. First, our empirical strategy focuses on social learning from chain establishments to community establishments. We recognize, however, that community establishments' closure decisions might also have a reciprocal impact on the closure decisions of chain establishments. At the moment, we do not have a comparable instrumental variable strategy to help pin down this effect. Second, as rich as the location tracking data from SafeGraph are, they are ill-suited to pinning down the mechanisms that underlie the patterns we observe (e.g., the exact nature of social information that firms acquire from competitors).

In sum, this study documents that community establishments are prone to following suit in responding to the closure decisions of chain establishments. Understanding organizational decision making during tumultuous times such as a global pandemic can help local governments design more effective policies to positively influence the behavior of local establishments.

## References


Akbarpour, Mohammad, Cody Cook, Aude Marzuoli, Simon Mongey, Abhishek Nagaraj, Matteo Saccarola, Pietro Tebaldi, Shoshana Vasserman, and Hanbin Yang. 2020. "Socioeconomic Network Heterogeneity and Pandemic Policy Response." *University of Chicago, Becker Friedman Institute for Economics Working Paper* .

Anderson, Roy M, Hans Heesterbeek, Don Klinkenberg, and T Déirdre Hollingsworth. 2020. "How Will Country-Based Mitigation Measures Influence the Course of the COVID-19 Epidemic?" *The Lancet* 395:931–934.

Azoulay, Pierre, Christopher Liu, and Toby E. Stuart. 2017. "Social Influence Given (Partially) Deliberate Matching: Career Imprints in the Creation of Academic Entrepreneurs." *American Journal of Sociology* 122:1223–1271.

Baker, Scott R, Nicholas Bloom, and Steven J Davis. 2016. "Measuring Economic Policy Uncertainty." *The Quarterly Journal of Economics* 131:1593–1636.

Balla-Elliott, Dylan, Zoë B Cullen, Edward L Glaeser, Michael Luca, and Christopher T Stanton. 2020. "Business Reopening Decisions and Demand Forecasts During the COVID-19 Pandemic." Technical report, National Bureau of Economic Research.

Benzell, Seth G, Avinash Collis, and Christos Nicolaides. 2020. "Rationing Social Contact During the COVID-19 Pandemic: Transmission Risk and Social Benefits of US Locations." *Proceedings of the National Academy of Sciences* .

Boxell, Levi, Jacob Conway, Matthew Gentzkow, Michael Thaler, and David Yang. 2020. "Polarization and Public Health: Partisan Differences in Social Distancing during the Coronavirus Pandemic." Technical report, National Bureau of Economic Research.

Chang, Serina, Emma Pierson, Pang Wei Koh, Jaline Gerardin, Beth Redbird, David Grusky, and Jure Leskovec. 2020. "Mobility network models of COVID-19 explain inequities and inform reopening." *Nature* pp. 1–8.

Chiou, Lesley and Catherine Tucker. 2020. "Social distancing, internet access and inequality." Technical report, National Bureau of Economic Research.

Coleman, James Samuel. 1990. *Foundations of Social Theory*. Belknap Press of Harvard University Press.

Davis, Gerald F. 1991. "Agents without principles? The spread of the poison pill through the intercorporate network." *Administrative science quarterly* pp. 583–613.

Derenoncourt, Ellora. 2019. "Can you move to opportunity? Evidence from the Great Migration." *Working Paper* .

DiMaggio, Paul and Filiz Garip. 2012. "Network effects and social inequality." *Annual review of sociology* 38:93–118.

DiMaggio, Paul J. and Walter W. Powell. 1983. "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields." *American Sociological Review* 48:147–160.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift. 2018. "Bartik instruments: What, when, why, and how." Technical report, National Bureau of Economic Research.

Goolsbee, Austan, Nicole Bei Luo, Roxanne Nesbitt, and Chad Syverson. 2020. "COVID-19 Lockdown Policies at the State and Local Level." *University of Chicago, Becker Friedman Institute for Economics Working Paper* .

Kelly, Erin L. 2003. "The strange history of employer-sponsored child care: Interested actors, uncertainty, and the transformation of law in organizational fields." *American journal of Sociology* 109:606–649.

Lee, David L, Justin McCrary, Marcelo J Moreira, and Jack Porter. 2020. "Valid t-ratio Inference for IV." *arXiv preprint arXiv:2010.05058* .

Pfeffer, Jeffrey, Gerald R Salancik, and Huseyin Leblebici. 1976. "The Effect of Uncertainty on the Use of Social Influence in Organizational Decision Making." *Administrative Science Quarterly* 21:227–245.

Powell, Walter W, Kenneth W Koput, and Laurel Smith-Doerr. 1996. "Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology." *Administrative science quarterly* pp. 116–145.

Rogers, Everett M. 2010. *Diffusion of innovations*. Simon and Schuster.

Shalizi, Cosma Rohilla and Andrew C. Thomas. 2011. "Homophily and Contagion are Generically Confounded in Observational Social Network Studies." *Sociological Methods & Research* 40:211–239.

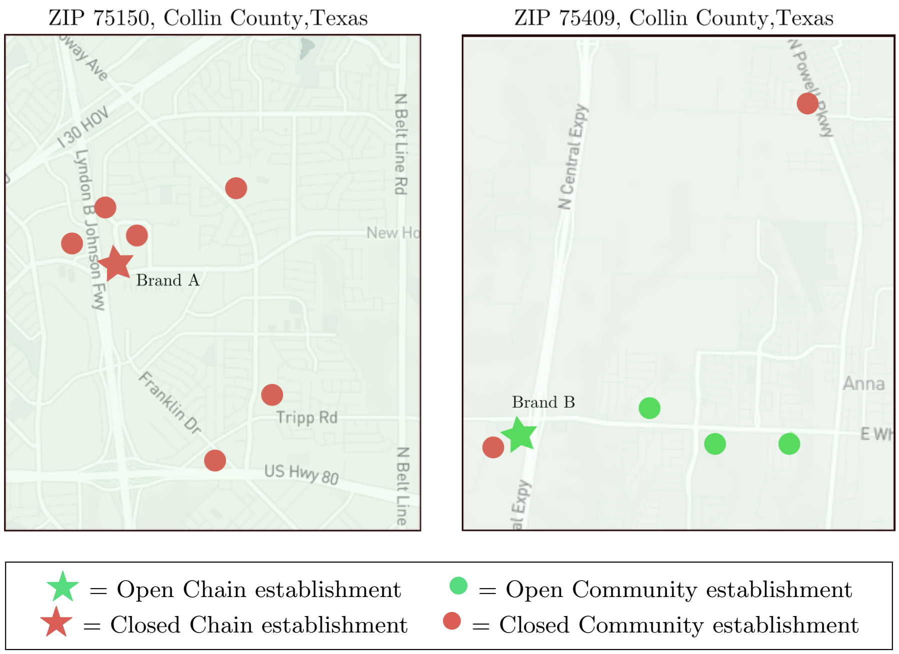
Turner, John C. 2010. "Social Categorization and the Self-concept: A Social Cognitive Theory of Group Behavior." In *Key readings in social psychology. Rediscovering social identity*, edited by T. Postmes and N.R. Branscombe, pp. 243–272. Psychology Press.

Tushman, Michael L and Elaine Romanelli. 1983. "Uncertainty, Social Location and Influence in Decision Making: A Sociometric Analysis." *Management Science* 29:12–23.

Young, H Peyton. 2009. "Innovation diffusion in heterogeneous populations: Contagion, social influence, and social learning." *American economic review* 99:1899–1924.

# Figures and Tables

Figure 1: Illustration of our research design




*Note:* This figure provides an illustration of our research design using two neighboring zip codes in Collin County, TX – 75150 on the left and 75409 on the right. This figure plots all establishments belonging to the NAICS code 713940 (Fitness and Recreational Sports Centers) in these two zip codes in our data. Chain establishments are represented by stars, while community establishments are indicated by circles. Establishments colored in red are closed, while establishments in blue are open on March 25, 2020 according to our algorithm. As is clear from this figure, the chain establishment in the left panel (which is Brand A) has closed while the chain establishment in the right panel (Brand B) has not. In this example, 78% of community establishments in the left panel are closed while only 66% of community establishments in the right panel are closed on this given day.

Table 1: Summary Statistics

| | N | Median | Mean | St. Dev. | Min | Max |
|---|---|---|---|---|---|---|
| Open | 10,368,135 | 0 | 0.431 | 0.495 | 0 | 1 |
| Prop. Branch Est. Open$_{t-1}$ | 10,368,135 | 0.7 | 0.593 | 0.356 | 0 | 1 |
| National Chain Opening Exposure$_{t-1}$ | 10,368,135 | 0.609 | 0.611 | 0.253 | 0.000 | 1.000 |
| Avg. February Traffic | 230,403 | 12.562 | 17.185 | 25.032 | 7.032 | 7,014.948 |
| Prop. Devices At Home | 9,688,964 | 0.343 | 0.342 | 0.110 | 0.003 | 0.778 |
| Shelter in Place | 10,368,135 | 0 | 0.156 | 0.362 | 0 | 1 |

Note: *Open* is a binary variable indicating whether the community establishment is open or not on a given day. *Prop. Branch Est. Open* is the ratio of chain establishments that remains open to the total number of chain establishments in the same industry and zip code. *National Chain Opening Exposure* is a weighted average of national opening rates of different chains that have establishments in the zip area and are in the same industry as the community establishment. We include three control variables. *Avg. February Traffic* captures the average number of visitors of an establishment in the month prior to the outbreak of COVID-19. *Prop. Devices At Home* measures the ratio of the total number of devices at home and the total number of devices in a census tract area. *Shelter In Place* is a binary variable indicating whether a Shelter-in-Place order is in effect in the community establishment's county. The *Local Customers*, *Loyal Customers*, and *Avg. February Traffic* are all cross-sectional measures measured in February. The other variables change over time, with *Prop. Devices At Home* missing some observations. If our models include this variable, we drop the missing cases.

Table 2: OLS and IV Regressions of Community Est. Open

**Panel A: OLS Estimates**

| | Open | | | |
|---|---|---|---|---|
| | Model 1 | Model 2 | Model 3 | Model 4 |
| Prop. Branch Est. $\text{Open}_{t-1}$ | 0.293*** | 0.180*** | 0.060*** | 0.037*** |
| | (0.011) | (0.006) | (0.003) | (0.002) |
| Avg. February Traffic | 0.001 | 0.001 | 0.001 | 0.001* |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| Prop. Devices At Home | −1.275*** | −1.993*** | −0.351*** | −0.168*** |
| | (0.068) | (0.040) | (0.027) | (0.026) |
| Shelter In Place | −0.042*** | −0.038*** | | |
| | (0.011) | (0.007) | | |
| (Intercept) | 0.678*** | | | |
| | (0.030) | | | |
| Fixed Effect NAICS | No | Yes (25) | Yes (25) | No |
| Fixed Effect Zip | No | Yes (11,879) | Yes (11,879) | Yes (11,879) |
| Fixed Effect Date | No | No | Yes (45) | No |
| Fixed Effect NAICS × Date | No | No | No | Yes (1,125) |
| Fixed Effect County × Date | No | No | No | Yes (70,738) |
| Observations | 9,688,964 | 9,688,964 | 9,688,964 | 9,688,964 |
| $\text{R}^2$ | 0.195 | 0.263 | 0.302 | 0.300 |

**Panel B: IV Estimates**

| | Model 5 | | Model 6 | |
|---|---|---|---|---|
| | First Stage | IV | First Stage | IV |
| Prop. Branch Est. $\text{Open}_{t-1}$ | | 0.113*** | | 0.099*** |
| | | (0.006) | | (0.008) |
| National Chain Opening $\text{Exposure}_{t-1}$ | 0.961*** | | 0.972*** | |
| | (0.013) | | (0.015) | |
| Avg. February Traffic | 0.000 | 0.001 | 0.000 | 0.001 |
| | (0.000) | (0.001) | (0.000) | (0.001) |
| Prop. Devices At Home | −0.232*** | −0.332*** | −0.004 | −0.165*** |
| | (0.032) | (0.026) | (0.024) | (0.024) |
| Fixed Effect NAICS | Yes (25) | Yes (25) | No | No |
| Fixed Effect Zip | Yes (11,879) | Yes (11,879) | Yes (11,879) | Yes (11,879) |
| Fixed Effect Date | Yes (45) | Yes (45) | No | No |
| Fixed Effect NAICS × Date | No | No | Yes (1,125) | Yes (1,125) |
| Fixed Effect County × Date | No | No | Yes (70,738 ) | Yes (70,738) |
| Observations | 9,688,964 | 9,688,964 | 9,688,964 | 9,688,964 |

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Note: Establishment-date level observations. The sample includes all establishment-dates from March 2 2020 to April 15 2020 for 230,403 community establishments who have at least one national brand competitor in their zipcode. The main dependent variable Open:0/1=1 if a place is deemed open according to our measure derived from SafeGraph cellphone visit data. Prop. Chain Est. Open is the percent of same-industry zip chain establishments that are open on the same date. Avg. February Foot Traffic denotes average visitors in February 2020. Prop. Devices at Home indicates number of devices that sheltered-in-place in a given zip code by not leaving their residence even once. Shelter In Place: 0/1=1 if a zip has a formal Shelter In Place regulation asking establishments to shut down. In the IV regression (Panel B), National Chain Opening Exposure indicates the predicted likelihood of competitor chain establishments being open as measured by national closing patterns. The number of fixed effects estimated are included in parentheses following the fixed effects indicators.