Rationing social contact during the COVID-19 pandemic: Transmission risk and social benefits of US locations

Seth G. Benzell[a,b,1,2], Avinash Collis[a,c,1,2], and Christos Nicolaides[a,d,1,2]

[a]MIT Initiative on the Digital Economy, Massachusetts Institute of Technology, Cambridge, MA 02142; [b]Argyros School of Business and Economics, Chapman University, Orange, CA 92866; [c]McCombs School of Business, The University of Texas at Austin, Austin, TX 78705; and [d]School of Economics and Management, University of Cyprus, 2109 Aglantzia, Nicosia, Cyprus

Edited by Paul R. Milgrom, Stanford University, Stanford, CA, and approved May 29, 2020 (received for review April 24, 2020)

To prevent the spread of coronavirus disease 2019 (COVID-19), some types of public spaces have been shut down while others remain open. These decisions constitute a judgment about the relative danger and benefits of those locations. Using mobility data from a large sample of smartphones, nationally representative consumer preference surveys, and economic statistics, we measure the relative transmission reduction benefit and social cost of closing 26 categories of US locations. Our categories include types of shops, entertainments, and service providers. We rank categories by their trade-off of social benefits and transmission risk via dominance across 13 dimensions of risk and importance and through composite indexes. We find that, from February to March 2020, there were larger declines in visits to locations that our measures indicate should be closed first.

COVID-19 | social contact | transmission risk | social welfare

Coronavirus disease 2019 (COVID-19) is primarily spread by droplets of mucous and saliva from those who are infected. Infected people are often asymptomatic (1), so, in the absence of a comprehensive system to test and trace individuals by infection status, all physical proximity is potentially dangerous. To address this concern, policy makers have implemented a wide variety of regulations on work, locations, and gatherings. Perhaps due to the infeasiblity of directly restricting visitor density, many of these restrictions vary by the type of the location.

We conceptualize the decision to shut down a location as a trade-off between infection risk and benefits. In this paper, we make an empirical contribution regarding which types of locations pose the best and worst risk–reward trade-offs. Governments should use this analysis to inform their decision making as they attempt to achieve their public health goals (such as $R < 1$) at minimum social cost. To do so, we combine several measures of the importance and danger of categories of stores and locations. We consider 26 categories that correspond to North American Industry Classification System (NAICS) industries or combinations thereof.

Danger and Importance Measures

We collect data of three types. These are data on the category's transmission risk, economic output and costs, and consumer value.

To quantify the potential contribution of a location to disease transmission (i.e., its danger), we utilize a fine-grained dataset on mobility from approximately 47 million smartphone devices in the United States. The data are collected by Safegraph, and record visits to 6 million points of interest. The "visitation" data include information about the total number of visits, total number of visitors, home census tract of visitors, and timing and length of visits. The "points of interest" data include information on location (full address), six-digit NAICS code, branding, and area.

The 26 location categories of interest account for $\sim$57% of all unique visits from January 2019 through March 2020. Out of all categories, full-service restaurants (sit-down) is the most popular in terms of both number of visits and unique visitors. Between February and March 2020, we observe a 24.9% drop in the total number of visits at all locations included in the analysis, reflecting the social distancing implemented in March. To account for the fact that our data cover only a fraction of individuals in the United States, we upscale every observed visit as a function of the visitor's home census tract to approximate the real number of visits and visitors for each location.

We create nine monthly level measures of a location's danger. Four are based directly on total visit data. These are total visits, total unique visitors, person-hours of visits during crowding of more than one visitor per 113 square feet (reflecting the Centers for Disease Control and Prevention's "six-foot" social distancing rule), and person-hours of visits during crowding of more than one visitor per 215 square feet (reflecting the German social distancing guideline of one customer per 20 m^2).

Using information on the home census tract of visitors, we add five additional measures. Four are analogous to the first four, but restricted to visits by individuals age 65 y and older (age estimates are based on the assumption that visits from older guests are proportional to their share of a census tract's population; we use visits from all guests in calculating location density). The final danger measure is the median distance traveled to a location.

To identify the cumulative danger of an entire category of locations, we sum the individual measures of all locations within the category (except for distance traveled, where we use the visit-weighted average). While, in this analysis, we weigh all nine danger measures equally, our results are very similar to those when restricting attention to the first four danger measures.

We measure the benefits of a location as coming from both its economic and consumer contributions. Our economic data come from the most recent edition of the US Census Statistics of US Businesses. Across our 26 categories, there are 1,427,433 firms and 2,024,839 establishments, compared to 2,029,514 geolocation points of interest. Our measures of economic importance consist of annual payroll, receipts, and employment. Our 26 categories encompass 32 million employees, 1.1 trillion dollars

Author contributions: S.G.B., A.C., and C.N. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no competing interest.

Data deposition: Data, survey instrument, derived statistics and code are available at GitHub (https://github.com/chrisnic12/RationingContact).

[1]To whom correspondence may be addressed. Email: benzell@chapman.edu, avinash.collis@mccombs.utexas.edu, or nicolaides.christos@ucy.ac.cy.

[2]S.G.B., A.C., and C.N. contributed equally to this work.

First published June 10, 2020.

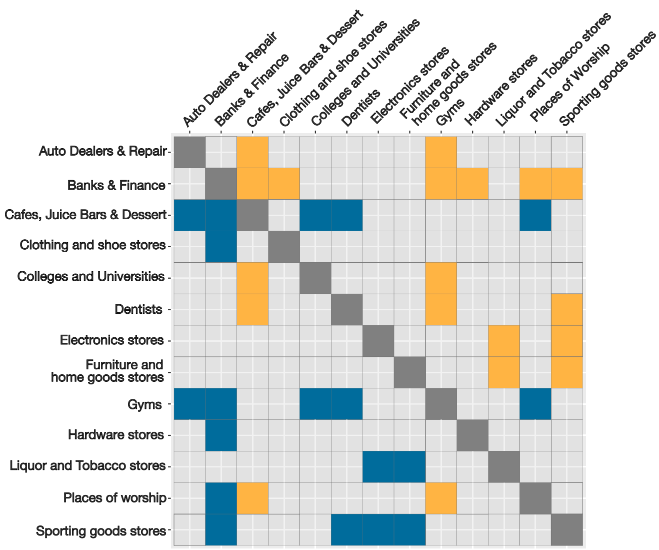

Fig. 1. Grid indicating dominating and dominated categories. A cell is gold if the row category is better on all nine risk and four importance dimensions than the column category, and blue for the converse.

in annual payroll, and 5.6 trillion dollars in annual receipts.[*] To measure consumer welfare, we conducted a nationally representative survey of 1,099 US residents. Respondents were recruited through Lucid, a market research firm, during April 13 to April 15, 2020. The survey was determined to be exempt by Massachusetts Institute of Technology's (MIT's) Institutional Review Board (Project E-2115). The sample is representative by age, gender, ethnicity, and region (2). Each respondent takes part in a series of single binary discrete choice experiments (3) where they choose which location, among two options, they would prefer to be open, whether or not the location is currently open. Discrete choice experiments have been widely used to measure valuations of market and nonmarket goods. To make responses consequential and incentivize respondents to respond truthfully, we gave them a chance to earn an additional monetary reward which was linked with their choices (4). Each respondent participated in a series of binary discrete choice experiments. We solicited a total of 32,970 decisions. Our consumer importance measure ranks categories by the proportion of times it is preferred over others.

Results

We juxtapose how different locations fare along our four dimensions of importance (consumer importance, employment, payroll, and receipts) and nine dimensions of transmission risk (visits, unique visitors, person hours at moderate density, and person hours at high density; the same four measures for only individuals age 65 y and older; and average median distance traveled). The core idea is that locations offering better trade-offs should face looser restrictions. The most conservative way to make this comparison is to look at whether there are any locations that dominate another in all dimensions of lower transmission danger and higher importance. This measure is conservative in the sense that any possible weighed aggregate measure of risk or importance will yield the same pairwise comparison.

Of our 26 categories, 13 do not dominate and are not dominated by any other. Of the 13 remaining categories, 1) gyms and 2) cafes, juice bars, and dessert parlors are the two categories with the most dominated pairings (Fig. 1). According to our measure, each of these locations should be opened only after banks, dentists, colleges, places of worship, and auto dealers and repair shops. Within types of stores, we find electronics stores and furniture stores should be opened before liquor and tobacco stores and sporting goods stores. The two locations that come out the best in this measure are banks and finance, with six dominant pairwise comparisons, and dentists, with three dominant pairwise comparisons.

Another way to determine which locations offer the best trade-offs is to create overall indexes of danger and importance, and to look for outliers. We create our danger index as the average rank of a category in the nine danger measures. We create our importance index as the average rank of a category in our three economic importance and one consumer importance measures. We up-weight the consumer importance measure so that it is equally weighted with the three economic importance measures.

There is a strong positive relationship between the danger of a category and its importance (Fig. 2*A*). However, there are clear outliers. Categories in the top left corner have high importance but low danger, and vice versa for categories in the bottom right corner. We estimate a linear regression, including an intercept term, of the importance index as a function of the danger index. Categories are colored by the value of the residual, which corresponds to the quality of the trade-off. We find that banks, general merchandise stores (e.g., Walmart), dentists, grocery stores, and colleges and universities should face relatively loose restrictions. Gyms, sporting goods stores, liquor and tobacco stores, bookstores, and cafes should face relatively tight restrictions. Our methodology also allows us to do similar analyses for different regions. Splitting our analysis by metropolitan and nonmetropolitan locations yields remarkably similar results, suggesting that the urban–rural divide is not an important dimension for policy makers.

There is a dramatic decrease in visits to all of these locations from February to March 2020. A natural final question is whether these reductions in visits are spread evenly across locations, or whether the reductions follow the risk–reward trade-off we measure. Fig. 2*B* plots the percent decrease in visits to a location type, from February to March 2020, as a function of "importance–risk trade-off favorability" (i.e., the gold to blue categorization in Fig. 2*A*). Weighing by February 2020 visits, there is a strong positive relationship. This suggests that at least some of the cost–benefit analysis we measure is being internalized by US consumers, businesses, and policy makers. Two of the largest outliers are 1) colleges and universities and 2) hardware stores. We find colleges to offer a relatively good trade-off, but most have shut down, leading to a 61% decline in visits. Conversely, we find liquor and tobacco stores to be relatively poor trade-offs (due to mediocre economic importance and small busy stores), yet the number of visits to this category has declined by less than 5%. Hardware stores are the location which has seen the largest increase in visits, as individuals scrounge for personal protective equipment and other home supplies. It is important to note that these visitation changes are due to a mix of federal, state, and local government, business, and individual level actions.

Discussion

A potential limitation of this analysis is that visitors to some location types are more concentrated within the space. We can partly account for this effect by measuring which locations offer services that require close physical proximity. In a complementary analysis, we merge in Occupational Employment Statistics

[*]Usually, public economic analyses of welfare exclude changes in labor costs in evaluating a policy, because the workers directly employed by the policy would have collected the same wage elsewhere. However, during this crisis, there is dramatic underemployment. Therefore, the work forces of these industries have very low opportunity costs, and their production should be counted in social surplus.

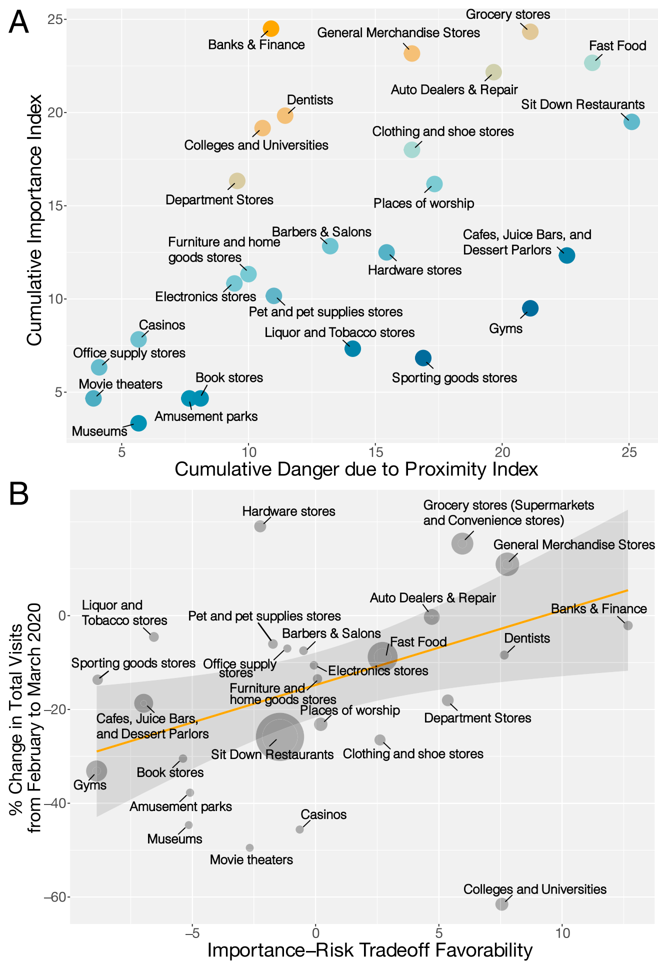

Fig. 2. (*A*) Category cumulative importance index and cumulative danger index. The color scale reflects the residuals, by category, of a linear regression of the importance index on the danger index. Golden categories have disproportionately high importance for their risk, and blue categories have disproportionately low importance. (*B*) Change in location category visits versus the category importance–risk residual. Marker sizes are proportional to total visits in February 2020.

(OES) data on occupational employment mix by category and Occupational Information Network (O*NET) data on occupational tasks. 1) Dentists and 2) barbershops and salons are the only two categories with a high share of workers requiring intense physical proximity (72% and 58%, respectively, of workers in these industries have proximity scores of over 90%). Additionally, movie theaters, gyms, and amusement parks have a high share of workers requiring a moderately high level of physical proximity (57%, 48%, 42%, respectively, of workers in these industries have proximity requirement scores of over 80%). We do not include these data in our main analysis because the need to be in close contact with visitors impacts both the risk of a category and the economic cost of shutting a category down. A category with a high share of workers who do not require close proximity to visitors will find it easier to reengineer itself to increase social distance, as well as to allow for work from home. Most retailers can offer curbside pickup rather than forcing customers to enter crowded stores (and, indeed, most states are encouraging curbside pickup). Locations can also be made safer through use of masks. This is especially important for locations like museums, with limited physical touching. On the other hand, locations like gyms both emphasize physical contact and make mask use unpleasant.

A more important limitation of our data is that we incorporate no information about linkages or complementarities between industries. If one industry is shut down, it could decrease the revenues, employment, consumer surplus, and visits of another (e.g., by depriving them of an important input), or increase them (e.g., by effectively "raising the cost" of a close substitute; we may be seeing this with restaurants and grocery stores). In the current analysis, we effectively assume that all industries are perfect substitutes.

There are other limitations in our analysis in terms of factors that impact our rankings of both location importance and risks. On the importance side, our binary choices do not yield information on the intensity of preferences, and leave out potentially important externalities from some locations (on mental or physical fitness, for example). Moreover, our survey sample size is limited, and further research should use a major survey research firm and larger samples. On the risk side, we fail to account for the fact that some locations encourage reckless physical activities or might disproportionately accommodate "superspreaders."

Governments and civic organizations across the world have made different decisions about how to implement and relax social distancing measures. As they do so, they have various tools at their disposal. In the United States, many of these restrictions have been location category specific. Details have varied from state to state, with gyms, places of worship, and liquor stores receiving particularly heterogeneous treatment. Why are different states adopting such different policies? One possibility is state-level variation in the importance or danger of a category. This variation would have to be separate from urban–rural heterogeneity, which we find to not make much of a difference.

Another possibility is that, in the absence of empirical evidence, states are being forced to make decisions in the dark. If so, we recommend that policy makers conduct analyses similar to the ones described in this paper, specific to their regions. Regional mobility, credit card transaction, and other relevant data are available from a variety of sources. This should be complemented with regularly conducted large-scale online consumer preference surveys to account for heterogeneity across regions, demographics, and time.

ACKNOWLEDGMENTS. We thank Jonathan Wolf and Safegraph for data. We thank Erik Brynjolfsson, Sinan Aral, and Dean Eckles for invaluable feedback. We additionally thank the MIT Initiative on the Digital Economy for research funding. We thank Victor Yifan Ye for help with visualizations. We thank Maxwell H. Levy, M.D. for medical insights. We thank Manuela Collis for feedback on survey design.

1. Y. Bai *et al.*, Presumed asymptomatic carrier transmission of covid-19. *JAMA* **323**, 1406–1407 (2020).
2. A. Coppock, O. A. McClellan, Validating the demographic, political, psychological, and experimental results obtained from a new source of online survey respondents. *Res. Polit.* **6**, 205316801882217 (2019).
3. R. T. Carson, T. Groves, Incentive and informational properties of preference questions. *Environ. Resour. Econ.* **37**, 181–210 (2007).
4. R. T. Carson, T. Groves, J. A. List, Consequentiality: A theoretical and experimental exploration of a single binary choice. *J. Assoc. Environ. Resour. Econ.* **1**, 171–207 (2014).