DISCUSSION PAPER SERIES

DP15050

HOW MUCH DOES COVID-19 INCREASE WITH MOBILITY? EVIDENCE FROM NEW YORK AND FOUR OTHER US CITIES

Stephen J. Redding, Edward L Glaeser and Caitlin Gorback

INTERNATIONAL TRADE AND REGIONAL ECONOMICS

ISSN 0265-8003

HOW MUCH DOES COVID-19 INCREASE WITH MOBILITY? EVIDENCE FROM NEW YORK AND FOUR OTHER US CITIES

Stephen J. Redding, Edward L Glaeser and Caitlin Gorback

Discussion Paper DP15050
Published 16 July 2020
Submitted 13 July 2020

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

HOW MUCH DOES COVID-19 INCREASE WITH MOBILITY? EVIDENCE FROM NEW YORK AND FOUR OTHER US CITIES

Abstract

How effective are restrictions on geographic mobility in limiting the spread of the COVID-19 pandemic? Using zip code data for Atlanta, Boston, Chicago, New York (NYC), and Philadelphia, we estimate that total COVID-19 cases per capita decrease on average by approximately 20 percent for every ten percentage point fall in mobility between February and May 2020. To address endogeneity concerns, we instrument for travel by the share of workers in remote work friendly occupations, and find a somewhat larger average decline of COVID-19 cases per capita of 27 percent. Using weekly data by zip code for NYC and a panel data specification including week and zip code fixed effects, we estimate a similar average decline of around 17 percent, which becomes larger when we measure mobility using NYC turnstile data rather than cellphone data. We find substantial heterogeneity across both space and over time, with stronger effects for NYC, Boston and Philadelphia than for Atlanta and Chicago, and the largest estimated coefficients for NYC in the early stages of the pandemic.

JEL Classification: H12, I12, J17, R41

Keywords: COVID-19, mobility, Cities

Stephen J. Redding - reddings@princeton.edu
Princeton University and CEPR

Edward L Glaeser - eglaeser@harvard.edu
Harvard University

Caitlin Gorback - caitlin.gorback@gmail.com
National Bureau of Economic Research

Acknowledgements
We thank Sam Patterson and Andrei Freund for excellent research assistance. Any errors, opinions and omissions are our own. First Draft: June 2020. The U.S. Department of Transportation funded this project through a grant to NBER Project on Transportation in the 21st Century. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

1 Introduction

A central challenge in evaluating lock-downs and other restrictions on mobility in response to COVID-19 is estimating their effectiveness in limiting the disease's spread. This estimation is challenging for several reasons. Mobility restrictions are introduced as a response to disease outbreaks, individuals make mobility decisions based on the threat of infection, and the relationship between transmission and mobility depends on the composition of susceptible, infected and recovered agents.

To address these challenges, we combine weekly data on COVID-19 cases by zip code in New York City (NYC) and cross-sectional data for four other U.S. cities, information on mobility from SafeGraph cellular phone data and subway turnstile data for NYC, and exogenous variation in mobility from the ability to work remotely and designation as an essential worker in state shutdown orders. In our preferred instrumental variables specifications, we estimate that a ten percentage point (10pp) decrease in mobility leads to a 17-27 percent fall in COVID-19 cases per capita. We find substantial heterogeneity across both space and over time, with stronger effects for NYC, Boston and Philadelphia than for Atlanta and Chicago, and the largest estimated coefficients for NYC in the early stages of the pandemic.

The hypothesis that movement spreads COVID-19 inspired the stay-at-home orders adopted across the world in 2020. While any contagious disease can be propagated through human interaction, the actual link between mobility and contagion is mediated by the nature of the disease and the behavior of the travelers. The social benefits of regulations limiting mobility depend on the empirical magnitude of the link between mobility and disease. The link between mobility and contagion could be minimal if infections occurred mostly through intimate contact, as with sexually transmitted diseases, or large if dense transit hubs enabled super-spreading events.

We focus on the relationship between the logarithm of the COVID-19 cases per capita and the fall in mobility, relative to February 2020 or the same date one year earlier. We cannot determine if the disease spread through travel itself or through interactions at a final destination. We lack individual-specific COVID-19 tests, and consequently measure the prevalence of the disease by zip code.

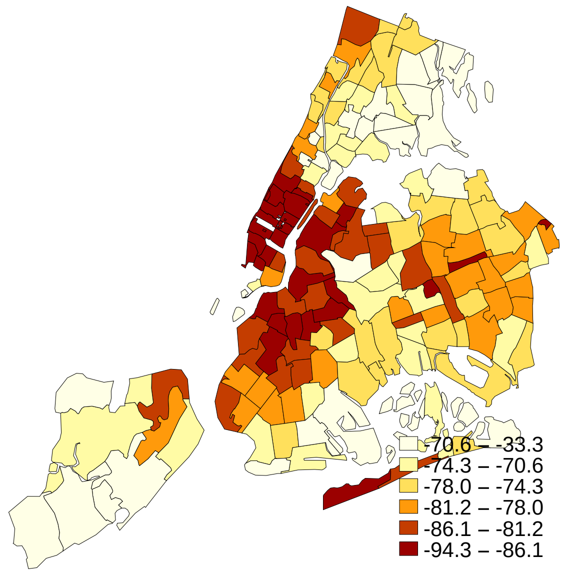
Neighborhood specific COVID-19 rate may not capture the true prevalence of the disease, because of differences in testing rates. We control for area-specific demographic variables that might predict testing rates, and for zip code fixed effects in our panel specifications. We also replicate our cross-sectional results on case rates with COVID-19 death rates in New York City. Unfortunately, the New York City zip code death data begins the week of May 18-24, 2020, and so cannot be used in our panel specifications.

Our primary mobility data source is the SafeGraph cellphone location data, which is available at the census blockgroup level nationwide. We aggregate travel data to the zip code to conform with our COVID-19 case data. For New York City, we supplement this data with zip code level information on turnstile use for public transit provided by the Metropolitan Transit Authority (MTA). For panel analysis, we aggregate the daily case and turnstile data to the week levels so that it is compatible with the weekly Safegraph data. As the impact of mobility on infection could change both with the level of infection in the population of travelers and the level of precaution, we test whether that impact changes over time in New York City. We split our sample into an early period that ends on April 26 and a later period, to test whether the impact of mobility on contagion was higher when cases where increasing or higher after the wave of infection peaked.

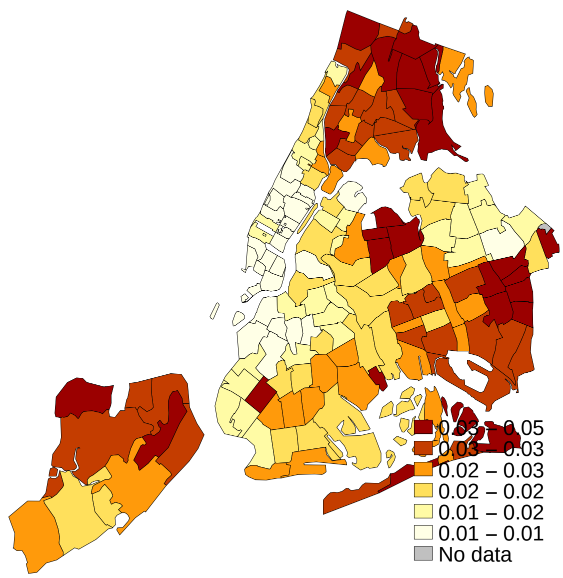
Figure 1 shows two maps of New York City that illustrate our core findings. The upper map shows the change in cell phone-measured mobility. The lower map shows the total COVID-19 cases per person as of June 1, 2020. In the parts of New York where mobility fell, case rates have been low. In the areas of New York where mobility remained higher, COVID-19 cases per person are higher. Appendix Figure A1 shows the correlation of 0.44 between change in the number of trips and the number of COVID-19 cases per capita.

This cross-sectional relationship suggests an elasticity of cases with respect to mobility of approximately two, and we estimate a similar elasticity of deaths with respect to mobility. Yet there are many reasons to be skeptical about this estimate. First, New York zip codes differ along many dimensions, such as income and race, which may also influence the spread and measurement of COVID-19. Second, mobility itself may decline with the level of infection, which could bias downwards the estimated link between cases and mobility. Third, the connection between mobility and disease can differ across cities, both due to different initial

Figure 1: Mobility Change and COVID-19 Cases per Person in NYC



(a) % Change in Trips, May '20 vs May '19



(b) Cases per Person

Source: Cases per person from NYC Health Department, available at `https://www1.nyc.gov/site/doh/covid/covid-19-data.page`. % Change in trips from SafeGraph Weekly Patterns Data, using visitors traveling from home. % Change in trips calculated between May 13-19, 2019 and May 4-10, 2020 infection rates and because travel may take different forms.

We take two strategies to address omitted neighborhood characteristics. First, we control for racial composition, income and age, which shrink the measured connection between mobility and disease prevalence, so that a 10pp in trips is associated with a 10% decline in case rates in New York City. Across all five cities, the coefficient is somewhat smaller but still statistically significant.

Second, we look at results over time within New York City zip codes controlling for neighborhood fixed effects. We follow the medical literature (Lauer et al. (2020)) and estimate a model with a two-week gap between new cases and mobility, as measured by cell phone records and turnstile data. The average onset time is closer to 1 week, but this two-week gap should capture over 97.5% of cases. Controlling for week and zip code fixed effects, the link between COVID-19 prevalence and the turnstile measure remains significant and positive, but the correlation between cases and cell phone mobility disappears. A 10pp drop in public transportation use is associated with a 0.3 log point fall in COVID-19 cases per capita.

If movement falls more in places with more disease, then these fixed effect estimates underestimate the true link between contagion and mobility. Consequently, our preferred specifications follow an instrumental variables strategy that uses employment by industry in a given zip code to predict changes in mobility. Following Bartik et al. (2020)), we focus on the share of residents working in essential industries, according to state shutdown essential worker designations, or that can work remotely, according to Dingel and Neiman (2020). Bartik et al. (2020) confirm that Dingel and Neiman (2020)'s predictions about remote work during the pandemic have largely born out across industries. Locations with more essential workers have more travel, and locations with teleworkable residents have less travel. In our NYC panel specifications, we allow the instruments to have a different impacts week-by-week.

Across almost every specification, the measured link between mobility and disease is larger in these instrumental variable specifications, which suggests that the ordinary least squares estimates are biased downwards because of reverse causality. In our preferred multi-city specification with demographic controls, we estimate that a 10pp drop in travel is associated with a 0.27 log point drop in per capita COVID-19 prevalence. A 0.27 log point fall in COVID-19 represents 5 fewer cases per 1,000 inhabitants, from a sample mean of 17 per

1,000.

City-specific estimates produce higher coefficients in New York, Boston and Philadelphia and lower coefficients in Atlanta and Chicago. Mobility seems to spread COVID-19 in the northeastern cities, but not in other cities. This difference likely reflects the initial infection rate rather than the nature of mobility, since public transportation is also used in Chicago.

Zip codes across cities may still have significant unobserved heterogeneity driving disease spread, motivating our panel research design. Moving from the cross-sectional analysis to our NYC-Safegraph data panel, we estimate an instrumental variables coefficient of 0.016 with zip code fixed effects. The effect is larger when controlling for zip code characteristics rather than the fixed effects. In the NYC-MTA panel, we estimate a larger coefficient of 0.034 with zip code fixed effects, which is an order of magnitude larger than the ordinary least squares coefficients, and in line with the cross-sectional IV results. The turnstile results supports the view that cases rise with mobility, at least over this period in NYC, but does not provide a clean estimate of the impact of public transportation use on the spread of COVID-19 hypothesized by Harris (2020), because our instruments are not public transit specific.

Our paper is related to the broader emerging body of research on COVID-19 in economics. First, a macroeconomics literature has used Susceptible-Infected-Recovered (SIR) models to simulate the impact of policies such as lock-downs on disease burden and economic outcomes, including Acemoglu et al. (2020), Alvarez, Argente and Lippi (2020), Atkeson (2020) and Fernández-Villaverde and Jones (2020). Second, others have analyzed the spatial diffusion of COVID-19, including Antràs, Redding and Rossi-Hansberg (2020), Argente, Hsieh and Lee (2020), Birge, Candogan and Feng (2020), Fajgelbaum et al. (2020), Bisin and Moro (2020) and Cuñat and Zymek (2020). A third line of work has examined how agents' behavioral responses (e.g. social distancing) likely effect the dynamics of COVID-19, including Fenichel et al. (2011), Alfaro et al. (2020), Farboodi, Jarosch and Shimer (2020), and Toxveard (2020). Fourth, a more microeconometric literature has examined locations' observable characteristics within cities and across U.S. counties that correlate with COVID-19 incidence, including Almagro and Orane-Hutchinson (2020), Couture et al. (2020) and Desmet and Wacziarg (2020). Finally, other research has compared the spatial diffusion and economic impact of COVID-19 to previous pandemics such as the 1918 influenza, as in Barro, Ursúa and Weng (2020) and Correia, Luck and Verner (2020).

Section 2 discusses our data sources and Section 3 introduces our empirical strategies. Section 4 discusses the results found using the cross-section of zip codes in five cities. Section 5 discusses our panel results. Both Sections 4 and 5 include results using instrumental variables. Section 6 concludes.

2 Data

We build a weekly panel of zip codes for NYC, and take a cross-sectional snapshot of four other US cities: Atlanta, GA, Boston, MA, Chicago, IL and Philadelphia, PA. All of these cities provide new **case counts by zip code**. Counts of daily new cases and cumulative cases come from each city's (or county's) department of public health. We have daily case data for New York City from April 4, 2020 through June 7, 2020. Because this misses much of the run-up in cases, we set cumulative and new cases to 0 in 2020w11, and assume cases double weekly until 2020w14, for which we have data. We use snapshots of cumulative cases for the remaining cities. Atlanta has data as of June 2, 2020; Boston as of May 24, 2020; Chicago and Philadelphia as of June 6, 2020.

SafeGraph has released publicly available data for cell phone trips between December 31st, 2018 - present. We pull weekly data for our five cities. The data tracks the number of visitors to a point of interest (POI) in a given week. Every POI observation contains information on its census blockgroup, as well as the number of visitors by their home blockgroup. We construct an origin-destination (OD) matrix from these observed trips, assuming travel from home, by counting how many visitors travel from a specific origin blockgroup to each POI blockgroup. The data only shows OD pairs with at least 4 visitors, so the data undercounts pairs with low travel volume, introducing measurment error. Finally, we aggregate the blockgroup level OD matrices to zip codes in line with our COVID-19 case data.

Our second mobility datasource comes from the Metropolitan Transit Authority's **turnstile data**. Turnstile entries are collected every few hours, for each unique turnstile. We map each subway turnstile to a zip code, and count the entries each week by zip code.

Our instruments use Dingel and Neiman (2020)'s **teleworkability shares** by 2-digit NAICS and definitions of **essential industries** (4-digit NAICS) from Delaware and Minnesota, in combination with zip code level employment data from the American Community Survey (ACS). Details for instrument construction follow in Section 3.1.

To analyse how mobility interacts with demographics, we collect **demographic data** at the Zip level on share African American, median age, and median income from the American Community Survey. To identify Zip level shares of teleworkable and essential workers, we use Zip level **employment by industry** from the ACS, which identifies the industries in which residentswork, the details of which are listed in Table A2.

Appendix Table A1 lists the summary statistics for each of our research designs. The first panel lists statistics for zip codes in all 5 cities in our cross-sectional analysis; the middle panel lists statistics for the zip codes included in our NYC SafeGraph panel; the bottom panel lists statistics for the zip codes which contain an MTA turnstile. Notably, all three panels show large drops in mobility, between 63% and 71% for the average zip code.

3 Research Design

In order to estimate the relationship between declines in mobility and reducing the spread of COVID-19, we implement the following:

$$ln(TotalCases_i^{PC}) = \alpha + \beta\%\Delta Trips_i + City_c + \varepsilon_i \quad (1)$$

$$ln(NewCases_{it}^{PC}) = \beta\%\Delta Trips_{i,t-2} + zip_i + week_t + \varepsilon_{it} \quad (2)$$

Equation (1) regresses log total cases per capita in zip code i on the $\%\Delta$ in mobility, measured by SafeGraph trips leaving residential zip code i. Equation (2) regresses log daily new cases in zip code i in week t on $\%\Delta$ in mobility, measured by SafeGraph trips leaving residential zip code i or by the number of turnstile turns in residential zip code i.

Because residents are likely to reduce trips in response to increases in cases of COVID-19, β in both equations is likely biased downwards. Additionally, trips and cases may be measured with error, further attenuating β. To alleviate concerns of bias, we build instruments using pre-period information on one's proclivity to travel during the pandemic.

3.1 Building the Instrument

We construct two instruments to allay concerns of bias and measurement error. Both instruments use the American Community Survey's zip code level data on residents' employment by industry. We know the share of employment in industry classifications listed in Appendix Table A2.

For our first instrument, we use data on essential industries from Minnesota and Delaware. These states designated a subset of the 4-digit NAICS codes as essential, allowing these industries to remain open. Within each 2-digit NAICS grouping in Appendix Table A2, we calculate the national share of employment designated as essential. Using the ACS zip level data, we know employment by 2-digit NAICS grouping. Using the national essential share in combination with the zip codel employment composition, we construct $ShareEssential_i$ for a zip code i as the employment-weighted average essential share. For example, consider a zip code i with 100 residents working in two industries: 40 in NAICS 42, and 60 in NAICS 31-33. If 50% of the national employment in NAICS 42 is designated as essential, and 30% of NAICS 31-33 is designated essential, we construct $ShareEssential_i = \frac{0.5*40+0.3*60}{100} = 0.38$.

For our second instrument, we use Dingel and Neiman (2020)'s definition of teleworkable industries. They provide a list of 2-digit NAICS industry codes, along with the share of that industry that can reliably telecommute. Since the ACS data combines many of Dingel & Neiman's NAICS codes, we take simple averages across the sub-categories that we combine. Using the ACS zip level data on residents' employment by industry, we can calculate a zip code's share of workers who can reliably telecommute. As with $ShareEssential_i$, we take the employment-weighted average telecommuting share across industries within a zip to construct $ShareTelework_i$.

The relevance criterion requires that the share of teleworking or essential workers is correlated with the change in travel within a given zip code. Appendix Figure A2 shows that trips dropped more in zip codes with lower shares of essential workers or in those that could reliably telecommute. The exclusion restriction requires that the share of essential workers or telecommuters in 2018 does not impact COVID-19 cases except through taking trips from home.

3.2 Multicity IV

For the cross-sectional, multiple city specification, we track the log of cumulative COVID-19 cases per capita by zip code as of the dates in Section 2 and regress it on the $\%\Delta$ in travel, in percentage points, between May 2019 and May 2020. This yields a coefficient in the flavor of an elasticity: $100 \times \beta$ can be interpreted as the % increase in total cases per capita due to an additional percentage point change in trips. We instrument for mobility with both instruments, in the first stage shown in Equation (1.1). This provides variation in $\%\Delta Trips_i$ using the pre-COVID-19 employment mix, allowing us to estimate the second stage, Equation (1.2), without concern of behavioral responses such as staying home.

$$\%\Delta Trips_i = \delta + \gamma IV_i + City_c + \eta_t \tag{1.1}$$

$$ln(TotalCases_i^{PC}) = \alpha + \beta\%\Delta\widehat{Trips}_i + City_c + \varepsilon_i \tag{1.2}$$

3.3 Panel Design

To use the panel data in NYC, we begin by considering Equation (2). $100 \times \beta$ can be interpreted as the % increase in new cases per capita in one's home zip code associated with an additional 1pp increase in trips originating in the same zip code.[1] Once we instrument for trips, the design changes from Equation (2) to Equations (2.1) and (2.2):

$$\%\Delta Trips_{i,t-2} = \gamma IV_i \times week_t + zip_i + week_t + \eta_{it} \tag{2.1}$$

$$ln(NewCases_{it}^{PC}) = \beta\%\Delta\widehat{Trips}_i + zip_i + week_t + \varepsilon_{it} \tag{2.2}$$

The first stage regresses trip change relative to travel in 2020w9 two weeks ago on the

[1] We use the approximation of ln(x+1) when cases or new cases equals 0. Results are robust to using inverse hyperbolic sine.

intrument, which we interact with a week dummy to introduce temporal variation. We include zip code fixed effects to control for time-invariant characteristics. Week fixed effects control for city-level changes in virus awareness, shut-down, orders, etc. that would impact all locations.

4 Cross-Sectional Results

Table 1 shows our results using a cross section of 448 zip codes in Atlanta, Boston, Chicago, New York, and Philadelphia. Our core specification is to regress the logarithm of cases identified as of our city-specific snapshot date (Atlanta - June 2, 2020; Boston - May 24, 2020; Chicago and Philadelphia - June 6, 2020; NYC -June 7, 2020), on the percent change in trips between the week of May 13-19, 2019 and May 4-10, 2020. Table 2 shows the city-specifc results.

Table 1 column (1) shows the ordinary least squares coefficient where total cases per capita is regressed on the decline in mobility. The estimated coefficient is 0.0186, which implies that for every ten percentage points that travel fell between May of 2019 and 2020, the number of cases per capita falls by 0.19 log points. This specification includes metropolitan area fixed effects, but not zip code fixed effects. The average zip code reported 17 cases per 1,000 people, so a 10pp reduction in travel would lower this to 13.8 per 1,000.

In regressions (2)-(4) we include our three primary controls separately, and in the fifth regression, we include all three controls together. These controls will be used in other specifications throughout this paper, but we only directly report the coefficients here. In the specifications including the controls separately, each control is significant. Column (2) shows that a 10pp increase in percent African-American is associated with a 0.058 log point increase in the COVID-19 case rate; this gap between African-American and white case rates is a widely known fact (Yancy (2020)).

The coefficient on age in the third regression is strongly negative. This is in line with older people taking protective steps to avoid contagion such as staying home, because they face higher mortality risk. Column (4) documents the stark relationship between income and COVID-19 cases; we estimate an elasticity of the case rate with respect to income of -0.62.

Table 1: Multiple City Demographics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	$ln(Cases_i)$	$\ln(Cases_i)$	$\ln(Cases_i)$	$\ln(Cases_i)$	$\ln(Cases_i)$	$\ln(Cases_i)$	$\ln(Cases_i)$
	OLS	OLS	OLS	OLS	OLS	IV	IV
$\%\Delta Trips_i$	0.0186***	0.0152***	0.0187***	0.00809***	0.00715***	0.0474***	0.0282***
	(0.00242)	(0.00240)	(0.00231)	(0.00237)	(0.00250)	(0.00395)	(0.00720)
$\%AfAm_i$		0.577***			0.139*		0.133
		(0.0725)			(0.0804)		(0.0889)
$ln(Age_i)$			-0.800***		0.294		-0.211
			(0.212)		(0.227)		(0.291)
$ln(Inc_i)$				-0.621***	-0.627***		-0.331***
				(0.0528)	(0.0720)		(0.124)
R-Sq.	0.529	0.571	0.549	0.648	0.652		
Root MSE						0.604	0.483
Obs.	448	448	448	448	448	448	448
F-Stat.						98.15	86.69
				Fixed Effects			
CBSA	X	X	X	X	X	X	X

Notes: The dependent variable is total cases per capita in zip code i. Columns (1)-(5) implement versions of Equation (1), $ln(TotalCases_i) = \alpha + \beta\%\Delta Trips_i + City_c + \varepsilon_i$, each column adding additional demographics. Columns (6)-(7) implement versions of Equation (1.2), $ln(TotalCases_i) = \alpha + \beta\%\widehat{\Delta Trips}_i + City_c + \varepsilon_i$. Equation (1.1) available upon request. Specifications (6) and (7) include both the teleworking and essential share instruments. Robust standard errors included in parentheses.

Significance: $^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$

This coefficient is stable when we include all three variables together, but the other variables either flip sign or lose significance. In this cross-sectional specification income, rather than race or age, is the larger determinant of COVID-19 case rates.

One natural explanation for why income reduces COVID-19 rates is that richer people are better able to adjust their lives to avoid contagion. Reduced mobility is one margin of that adjustment, and richer areas have dramatically less mobility as of May 2020 in these cities. A 0.1 log point increase in the median income of the zip code is associated with a 1.2 percent drop in trips relative to May of last year. Yet despite controlling for the fall in trips, income remains an important explanatory variable, suggesting our trips variable captures only one dimension of protective behavior.

The coefficient on mobility remains stable when we control for either race or age, but the estimate halves when we include income, either on its own or as one of three control variables. In column (5), a 10pp reduction in mobility is associated with a 0.07 log point reduction in cases per capita.

Addressing the possibility that this coefficient is biased downwards because mobility falls more where COVID-19 cases spike, we now use our two instruments for mobility, as in Equations 1.1 and 1.2. The coefficient on mobility becomes much larger, both with and without controls, in regressions (6) and (7). The coefficient on income shrinks accordingly, as workers in teleworkable industries have on average higher incomes.[2]

The coefficient in regression (7) implies that as mobility drops by 10pp, COVID-19 case rates drop by 0.28 log points. The average zip code saw 17 cases per 1,000 people, so a 10pp drop in mobility would drop the case rate to 12.2 per 1,000. This yields 1.6 fewer cases per 1,000 than implied by the OLS analysis, for the same drop in travel. We take this as evidence supporting the view that cases have been much lower in places where workers could switch to remote working, but we are cautious about interpreting the coefficient as a causal estimate on trips alone. Workers in essential industries or industries that cannot be done remotely face risks from many places, especially infections in the workplaces. We therefore interpret this as suggestive evidence that remaining at home reduces COVID-19 exposure,

[2] See for example the American Time Use Survey, as discussed in `https://siepr.stanford.edu/research/publications/how-working-home-works-out`.

but not that we can identify COVID-19 infection rates with any particular act of mobility.

In Table 2, we break out our results for all five cities in our sample. We also include the results for New York City death rates in the first column. Panel A shows the ordinary least squares coefficient with no other controls. Panel B shows the ordinary least squares coefficient with the demographcis controls. Panel C shows the instrumental variables coefficient with controls.

In Panel A column (1), we show the 0.019 coefficient when the logarithm of death rates are regressed on mobility across New York City zip codes. This coefficient becomes insignificant and negative in the first column of Panel B, which adds controls. Death rates are strongly associated with age at the zip code level and with the share of the population that is African-American. In Panel C, we find that after instrumenting for mobility, the coefficient rises to 0.027, a 40% increase from Panel A. While the ordinary least squares coefficient on mobility does not survive controls, the instrumental variables coefficient is robust, reflecting the fact that deaths were much higher in those parts of New York where workers could not switch to remote work, or where essential workers live.

In the second column, we show our results for COVID-19 cases in New York City. The estimate in Panel A is 0.023, which is close to the deaths coefficient in the first column and the coefficient for all cities together in Table 1. When we control for demographics in Panel B, the coefficient falls by over fifty percent but remains significant. The instrumental variables coefficient is quite large. This may be due to downward bias in the ordinary least squares coefficients as mobility shrank in response to local outbreaks of COVID-19, the reason we seek an instrument, or because the instrument is correlated with the error term.

Column (5) shows results for Chicago, the other large city in our sample with more than 50 zip codes. The Chicago coefficient is comparable to the coefficient in New York City when we have no other controls. With controls, the coefficient for Chicago becomes small and statistically insignificant. The instrumental variables strategy does not change that fact for Chicago. These results suggest that mobility was less harmful in Chicago than it was in New York.

The other three cities have small samples of zip codes and we are wary of inferring much from their results. Philadelphia shows a coefficient of 0.011 with uncontrolled ordinary least

Table 2: Multiple City Results

	(1) ln($Deaths_i$) NYC	(2) ln($Cases_i$) NYC	(3) ln($Cases_i$) Atlanta	(4) ln($Cases_i$) Boston	(5) ln($Cases_i$) Chicago	(6) ln($Cases_i$) Philadelphia
			Panel A: OLS			
$\%\Delta Trips_i$	0.0187***	0.0229***	-0.00431	0.0510***	0.0186***	0.0114***
	(0.00615)	(0.00384)	(0.0197)	(0.0123)	(0.00317)	(0.00352)
R-Sq.	0.0902	0.219	0.00246	0.495	0.164	0.201
Obs.	159	159	22	19	206	42
			Panel B: OLS With Demographics			
$\%\Delta Trips_i$	-0.00466	0.00963***	-0.00402	0.0362*	0.00425	0.00702
	(0.00434)	(0.00345)	(0.0128)	(0.0192)	(0.00289)	(0.00503)
R-Sq.	0.476	0.436	0.549	0.532	0.512	0.436
Obs.	159	159	22	19	206	42
			Panel C: IV With Demographics			
$\%\widehat{\Delta Trips}_i$	0.0269**	0.0605***	0.00350	0.0668**	-0.00958	0.0131*
	(0.0117)	(0.0126)	(0.0329)	(0.0274)	(0.00995)	(0.00710)
Root MSE	0.405	0.443	0.563	0.439	0.483	0.226
Obs.	159	159	22	19	206	42
F-Stat.	40.36	35.39	4.480	4.512	47.30	4.310
			Controls for Panels B & C			
$\%AfAm_i$	X	X	X	X	X	X
$ln(Age_i)$	X	X	X	X	X	X
$ln(Inc_i)$	X	X	X	X	X	X

Notes: The dependent variable is total cases, or total deaths, per capita in zip code i. Panels A and B show versions of OLS Equations (1) for separate cities' cases, as well as NYC's deaths. Panel C shows results from Equation (1.2), adding additional demographic controls, X_i: $ln(TotalCases_i) = \alpha + \beta\%\widehat{\Delta Trips}_i + \Gamma X_i + \varepsilon_i$. Panel C uses both the telework and essential share instuments. Robust standard errors in parentheses. Significance: $^{*}p < 0.10, ^{**} p < 0.05, ^{***} p < 0.01$

squares and .013 with the instrumental variables results with controls. The ordinary least squares results with controls produce a considerably smaller coefficient. The Boston mobility coefficients are large and significant in all three specification. The Atlanta results are small and insignificant in all three specifications, possibly suggesting that mobility was not strongly associated with the spread of COVID-19 in Atlanta during this time period.

We believe that the results for the east coast cities tell a consistent story. In Boston, New York and Philadelphia, the coefficients in Panels A and C are statistically significant and sizable in magnitude. Mobility appears to have been reliably correlated with the spread of the pandemic in those cities. In Atlanta and Chicago, the correlation between COVID-19 and mobility is weak or non-existent. This pattern of results is consistent with the idea that the impact of mobility is related to the initial infection rate, which is likely to have been higher on the east coast.[3] It could also be that the east coast is more connected or shared transport is more prevalent there.

5 NYC Panel Results

We now turn to our panel results looking within New York City over time. We have better coverage for New York than elsewhere and match the number of new COVID cases with mobility using the Safegraph data in Table 3. In Table 4, we repeat those specifications using the MTA turnstile data. In both tables, the Panel A shows results for the entire sample. Panel B shows results splitting the sample in two halves: the first half of the sample as new cases were growing, and the second half of the sample, when new cases were falling.

Table 3 column (1) shows our ordinary least squares coefficient, with zip code and week fixed effects. Panel A shows that over the entire time period, there is no correlation between mobility and COVID-19 cases within zip code. This reflects the fact that the zip codes with the large drops in mobility did not necessarily experience fewer cases. Panel B shows that there is next to no relationship during the first period. This coefficient drops during the second period , showing a significant negative coefficient on mobility. As we find it difficult

[3] In the conventional SIR model, the rate of new infections depends on the product of the fractions of infected and susceptibles (Kermack and McKendrick 1927).

to imagine how reduced mobility could have increased the spread of COVID-19, we interpret this to mean that mobility shut down in places where COVID-19 cases were errupting.

Those fears of reverse causality inspire the remaining regressions. Table 3 column (2) shows results with our two instruments and no other controls. The effect is striking. If this coefficient were correct, then a 10pp drop in trips is associated with a 0.87 log point decrease in COVID-19 cases. Panel B shows that the effect is stronger in the first half of the sample than in the second half of the sample. This specification is comparable to the cross-sectional results above with no other controls. In both cases, the correlation between our instruments and the demographic variables is surely biasing this coefficient upwards.

Column (3) includes our three demographic controls. As expected, the coefficient drops and is in line with our previous results, and is significant at the 1% level. A ten percentage point fall in trips is associated with a 0.29 log point decline in cases per capita. The estimated coefficient is about fifty percent large in the first half of the sample relative to the second half of the sample.

Column (4) includes fixed effects for the five boroughs of New York City. The coefficients increase relative to column (3) without these borough controls. Controlling for borough causes the estimate for income to fall, because boroughs are strongly correlated with income and appear to have an independent impact on cases. As the estimate for income falls, the estimate for mobility rises.

The fifth and sixth columns show our results including zip code fixed effects. These effects absorb all of the cross-sectional variation across the city, and as in column (1), they essentially cause the coefficient to drop to zero. We include results with both instruments, and for the telecommuting instrument alone. We are most comfortable with our results for the entire period shown in the top panel, but when we include both instruments the first stage F-statistic falls to a worrying 1.9. With only the telecommuting instrument, the F-statistic remains at 4.4, marginally better, with the point estimate significant at the 5% level.

In the fifth column, we estimate a coefficient of 0.01 and in the sixth column, we estimate a coefficient of 0.016, the only one to be statistically significant. Both are similar in magnitude and comparable to earlier results, suggesting an elasticity of cases with respect to trips of

Table 3: NYC Panel Results: SafeGraph Trips

	(1)	(2)	(3)	(4)	(5)	(6)
	$\ln(New_{it})$	$\ln(New_{it})$	$\ln(New_{it})$	$\ln(New_{it})$	$\ln(New_{it})$	$\ln(New_{it})$
	OLS	IV	IV	IV	IV	IV
	Panel A: Full Sample					
$\%\Delta Trips_{i,t-2}$	-0.00170	0.0869***	0.0288***	0.0381***	0.00996	0.0161**
	(0.00194)	(0.00524)	(0.00679)	(0.00590)	(0.00714)	(0.00769)
Root MSE	0.382	0.769	0.509	0.501	0.386	0.392
Observations	2045	2045	2045	2045	2045	2045
First Stage F-Stat.	2446.3	274.5	155.1	97.98	1.949	4.382
	Panel B: Split Sample (2020w11 - 2020w17 vs. 2020w18 - 2020w23)					
$\%\Delta Trips_{i,t-2}$	0.000389	0.114***	0.0424***	0.0619***	0.0317***	0.0236**
$\times 1^{st} Half$	(0.00201)	(0.0107)	(0.00908)	(0.00865)	(0.00884)	(0.00916)
$\%\Delta Trips_{i,t-2}$	-0.00427*	0.0716***	0.0239***	0.0277***	0.00273	0.0128
$\times 2^{nd} Half$	(0.00238)	(0.00592)	(0.00717)	(0.00647)	(0.00724)	(0.00801)
Root MSE	0.381	0.822	0.527	0.536	0.407	0.396
Observations	2045	2045	2045	2045	2045	2045
First Stage F-Stat.	2297.3	129.9	117.1	65.34	11.27	3.314
	Controls					
$\%AfricanAmerican_i$			X	X		
$ln(Age_i)$			X	X		
$ln(Inc_i)$			X	X		
	Fixed Effects					
Zip_i	X				X	X
$Borough_i$				X		
$Week_t$	X	X	X	X	X	X
	Instrument(s)					
$ShareTele_i \times Week_t$	X	X	X	X	X	X
$ShareEss_i \times Week_t$	X	X	X	X	X	

Notes: NYC panel results using SafeGraph trips from home. Dependent variable is log of new cases per capita in zip code i in week t. Panels A shows results for the full panel, reporting β from Equation (1) in the first column, with versions of Equation (2.2) in columns (2)-(6): $ln(NewCases_{it}) = \beta\%\widehat{\Delta Trips_i} + zip_i + week_t + \varepsilon_{it}$. Panel B splits the time period in half, and interacts the coefficient of interest with the two time periods, decomposing β into $\beta^{1^{st}Half}$, $\beta^{2^{nd}Half}$. Columns (2)–(5) use both the telework and essential share instuments. Robust standard errors in parentheses.
Significance: $^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$

around one.

Breaking the sample into halves, the coefficients are significant and positive in the first period. The coefficients for the second half are not significantly different from 0. One interpretation of these results is that the mobility drove COVID-19 contagion through the end of April. After residents spent significant time without leaving home, mobility no longer drove contagion.

In Table 4, we turn to our results using New York City turnstile data. It is tempting to view this as providing an independent measure of the impact of public transportation trips, as opposed to all types of mobility. That view is tenable with the ordinary least squares results, if those results are not biased by reverse causality. That view is not tenable with our instrumental variable results, because we use the same instruments used for Safegraph mobility. We believe that using the same instruments for different variables is reasonable, as both variables are imperfect attempts to measure mobility.

Column (1) shows our ordinary least squares results with zip code fixed effects. The coefficient is positive and statistically significant, but modest in magnitude over the entire sample, 1^{st} and 2^{nd} halves. A ten percentage point fall in public transit trips is associated with 0.035 log points fewer COVID-19 cases.

Columns (2)–(4) show instrumental variables results without zip code fixed effects that closely parallel those found in columns (2)–(4) of Table 3. Across the entire time period, the coefficients with no controls in column (2) are about 0.08, as in Table 3, and the coefficients in (3) and (4) around both around 0.04. The results are quite similar in both the first half and the second half of the sample period, and they are uniformly stronger than those using the Safegraph data. It could be that the Turnstiles data captures a riskier form of mobility, perhaps due to trip duration or shared mode, or that the zip codes with subway stations have more people travelling in them or different levels of infection.

In columns (5) and (6), we show results with zip code fixed effects. As in Table 3, we show results using both instruments and only the telecommuting instrument. In this case, the results are quite similar. The coefficients for the overall period are 0.032 and 0.034, and significant at the 1% level. These imply a quite large impact of reducing trips. A 10pp reduction in trips is associated with a 0.33 log point fall in COVID-19 cases.

Table 4: NYC Panel Results: MTA Turnstile Trips

	(1) $\ln(New_{it})$ OLS	(2) $\ln(New_{it})$ IV	(3) $\ln(New_{it})$ IV	(4) $\ln(New_{it})$ IV	(5) $\ln(New_{it})$ IV	(6) $\ln(New_{it})$ IV
	Panel A: Full Sample					
$\%\Delta Trips_{i,t-2}$	0.00353***	0.0748***	0.0462***	0.0417***	0.0315***	0.0339***
	(0.00121)	(0.00476)	(0.00659)	(0.00659)	(0.00736)	(0.00865)
Root MSE	0.417	0.828	0.623	0.579	0.482	0.493
Observations	1523	1523	1523	1523	1523	1523
First Stage F-Stat.	1417.2	247.0	109.3	54.87	18.36	15.40
	Panel B: Split Sample (2020w11 - 2020w17 vs. 2020w18 - 2020w23)					
$\%\Delta Trips_{i,t-2}$ $\times 1^{st} Half$	0.00288**	0.0645***	0.0430***	0.0383***	0.0316***	0.0353***
	(0.00120)	(0.00569)	(0.00654)	(0.00644)	(0.00734)	(0.00876)
$\%\Delta Trips_{i,t-2}$ $\times 2^{nd} Half$	0.00557**	0.0957***	0.0606***	0.0563***	0.0404***	0.0591***
	(0.00222)	(0.00694)	(0.00935)	(0.00912)	(0.0114)	(0.0152)
Root MSE	0.417	0.839	0.642	0.597	0.492	0.537
Observations	1523	1523	1523	1523	1523	1523
First Stage F-Stat.	1320.1	159.3	93.70	44.41	9.310	8.845
	Controls					
$\%AfricanAmerican_i$			X	X		
$ln(Age_i)$			X	X		
$ln(Inc_i)$			X	X		
	Fixed Effects					
Zip_i	X				X	X
$Borough_i$				X		
$Week_t$	X	X	X	X	X	X
	Instrument(s)					
$ShareTele_i \times Week_t$	X	X	X	X	X	X
$ShareEss_i \times Week_t$	X	X	X	X	X	

Notes: NYC panel results using MTA turnstile trips in a given residential zip code. Dependent variable is log of new cases per capita in zip code i in week t. Panels A shows results for the full panel, reporting β from Equation (1) in the first column, with versions of Equation (2.2) in columns (2)-(6): $ln(NewCases_{it}) = \beta\widehat{\%\Delta Trips_i} + zip_i + week_t + \varepsilon_{it}$. Panel B splits the time period in half, and interacts the coefficient of interest with the two time periods, decomposing β into $\beta^{1^{st}Half}$, $\beta^{2^{nd}Half}$. Columns (2)–(5) use both the telework and essential share instuments. Robust standard errors in parentheses.
Significance: $^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$

6 Conclusion

Research is at an early stage on the progress of COVID-19 across America. Yet we already have plausible sources of variation in the behavior of different parts of the population. Some industries comfortably worked from home. Others could not and have braved exposure to COVID-19 to earn a living. In this paper, we used variation in that industrial mix to estimate the impact that mobility had on COVID-19 case rates.

Our estimates were not uniform. The measured effects of mobility were larger in New York, Boston and Philadelphia. They were smaller in Atlanta and Chicago. Moving around New York appears to have been riskier in March and early April than in May. Nonetheless, our estimates paint a consistent picture that mobility led to more COVID-19 exposure. Moreover, almost all estimates imply an elasticity greater than one, so that a 10pp drop in trips lead to a 0.1 log point or more reduction in COVID-19 cases per capita.

We do not claim these large effects would hold in different settings or when people wear masks and gloves while traveling. We hope these results may help future cost-benefit analyses around lockdown policies, but no policy implications follow directly from them. They simply remind us that people whose jobs required them to leave their homes were more likely to get COVID-19, and – at least in New York City – they were more likely to die.

References

Acemoglu, Daron, Victor Chernozhukov, Iván Werning, and Michael D. Whinston. 2020. "Optimal Targeted Lockdowns in a Multi-Group SIR Model." National Bureau of Economic Research Working Paper 27102.

Alfaro, Laura, Ester Faia, Nora Lamersdorf, and Farzad Saidi. 2020. "Social Interactions in Pandemics: Fear, Altruism, and Reciprocity." National Bureau of Economic Research Working Paper 27134.

Almagro, Milena, and Angelo Orane-Hutchinson. 2020. "The Determinants of the Differential Exposure to COVID-19 in New York City and their Evolution Over Time." New York University Working Paper.

Alvarez, Fernando, David Argente, and Francesco Lippi. 2020. "A Simple Planning Problem for COVID-19 Lockdown." National Bureau of Economic Research Working Paper 26981.

Antràs, Pol, Stephen Redding, and Esteban Rossi-Hansberg. 2020. "Globalization and Pandemics." Harvard University Working Paper.

Argente, David O, Chang-Tai Hsieh, and Munseob Lee. 2020. "The Cost of Privacy: Welfare Effect of the Disclosure of COVID-19 Cases." National Bureau of Economic Research Working Paper 27220.

Atkeson, Andrew. 2020. "What Will Be the Economic Impact of COVID-19 in the US? Rough Estimates of Disease Scenarios." National Bureau of Economic Research Working Paper 26867.

Barro, Robert J, José F. Ursúa, and Joanna Weng. 2020. "The Coronavirus and the Great Influenza Pandemic: Lessons from the "Spanish Flu" for the Coronavirus's Potential Effects on Mortality and Economics Activity." Working Paper.

Bartik, Alexander W, Marianne Bertrand, Zoë B Cullen, Edward L Glaeser, Michael Luca, and Christopher T Stanton. 2020. "How Are Small Businesses Adjusting to COVID-19? Early Evidence from a Survey." National Bureau of Economic Research Working Paper 26989.

Birge, John R., Ozan Candogan, and Yiding Feng. 2020. "Controlling Epidemic Spread: Reducing Economic Losses with Targeted Closures." Becker Friedman Institute Working Paper 2020-57.

Bisin, Alberto, and Andrea Moro. 2020. "Learning Epidemiology by Doing: The Empirical Implications of a Spatial SIR Model with Behavioral Responses." University Library of Munich, Germany MPRA Paper 101059.

Correia, Sergio, Stephan Luck, and Emil Verner. 2020. "Pandemics Depress the Economy, Public Health Interventions Do Not: Evidence from the 1918 Flu." National Bureau of Economic Research Working Paper 3561560.

Couture, Victor, Jonathan I Dingel, Allison Green, Jessie Handbury, and Kevin R Williams. 2020. "Measuring Movement and Social Contact with Smartphone Data: a Real-time Application to COVID-19." University of Chicago Booth School of Business Working Paper.

Cuñat, Alejandro, and Robert Zymek. 2020. "The (Structural) Gravity of Epidemics." CESifo CESifo Working Paper Series 8295.

Desmet, Klaus, and Romain Wacziarg. 2020. "Understanding Spatial Variation in COVID-19 across the United States." National Bureau of Economic Research Working Paper 27329.

Dingel, Jonathan I, and Brent Neiman. 2020. "How Many Jobs Can be Done at Home?" National Bureau of Economic Research Working Paper 26948.

Fajgelbaum, Pablo, Amit Khandelwal, Wookun Kim, Cristiano Mantovani, and Edouard Schaal. 2020. "Optimal Lockdown in a Commuting Network." National Bureau of Economic Research Working Paper 27441.

Farboodi, Maryam, Gregor Jarosch, and Robert Shimer. 2020. "Internal and External Effects of Social Distancing in a Pandemic." National Bureau of Economic Research Working Paper 27059.

Fenichel, Eli, Carlos Castillo-ChÃ¡vez, Graziano Ceddia, Gerardo Chowell, Paula Parra, Graham Hickling, Garth Holloway, Richard Horan, Benjamin Morin, Charles Perrings, Michael Springborn, Leticia Velazquez, and Cristina Villalobos. 2011. "Adaptive Human Behavior in Epidemiological Models." *Proceedings of the National Academy of Sciences of the United States of America*, 108: 6306–11.

Fernández-Villaverde, Jesús, and Chad Jones. 2020. "Estimating and Simulating a SIRD Model of COVID-19 for Many Countries, States, and Cities." National Bureau of Economic Research Working Paper 27128.

Harris, Jeffrey E. 2020. "The Subways Seeded the Massive Coronavirus Epidemic in New York City." National Bureau of Economic Research Working Paper 27021.

Kermack, William Ogilvy, and A. G. McKendrick. 1927. "A Contribution to the Mathematical Theory of Epidemics, Part I." *Proceedings of the Royal Society of London, Series A*, 115(772): 700–721.

Lauer, S. A., K. H. Grantz, Q. Bi, F. K. Jones, Q. Zheng, H. R. Meredith, A. S. Azman, N. G. Reich, and J. Lessler. 2020. "The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application." *Annals of Internal Medicine*, 179(9): 577–582.

Toxveard, Flavio. 2020. "Equilibrium Social Distancing." Cambridge Working Papers in Economics, Insitute for New Economic Thinking Working Paper 2020-08.

Yancy, Clyde W. 2020. "COVID-19 and African Americans." *JAMA*, 323(19): 1891–1892.

A Appendix Tables & Figures

Table A1: Summary Statistics

Variable	Mean	St.Dev.
	Panel A: All 5 Cities	
$TotalCases_i$	758	710
$Trips_i^{preCOVID}$	2682	3110
$Trips_i^{COVID}$	803	1119
$\%\Delta Trips_i$	-70	13
$ShareTele_i$	0.41	0.04
$ShareEss_i$	0.71	0.02
Pop_i	41,387	22,083
Age_i	37	5
Inc_i	87,026	43,312
$\%AfricanAmerican_i$	0.24	0.28
Observations	448	
	Panel B: NYC SafeGraph Panel	
$NewCases_{it}$	69	108
$\%\Delta Trips_{i_t-2}$	-63	26
$ShareTele_i$	0.42	0.04
$ShareEss_i$	0.72	0.02
Pop_i	51,887.44	24,919
Age_i	38	5
Inc_i	82,318	46,052
$\%AfricanAmerican_i$	0.24	0.25
Observations	2045	
	Panel C: NYC Turnstile Panel	
$NewCases_{it}$	71	111
$\%\Delta Trips_{i,t-2}$	-70	31
$ShareTele_i$	0.43	0.05
$ShareEss_i$	0.72	0.02
Pop_i	54,946	25,559
Age_i	36	4
Inc_i	87,880	58,057
$\%AfricanAmerican_i$	0.21	0.22
Observations	1523	

Notes: Case data from specific cities' or counties' health departments as in Section 2. Trips pre- and during COVID-19 from SafeGraph. Share telecommute and share essential as in Sections 2 and 3.1. Population, age, income and share African American from 2018 ACS data. Panel A uses cross-sectional data for all zips in the 5 cities. Panel B uses all zip codes in NYC from 2020w11 - 2020w23. Panel C uses all zip codes with subway turnstiles in NYC from 2020w11 - 2020w23.

Table A2: Industries and Codes Available in Zip Level ACS Employment Data

ACS Indsutry Description	Associated NAICS Codes
Agriculture, forestry, fishing and hunting, and mining	11, 21
Transportation and warehousing, and utilities	22, 48-49
Construction	23
Manufacturing	31-33
Wholesale trade	42
Retail trade	44-45
Information	51
Finance and insurance, and real estate and rental and leasing	52, 53
Professional, scientific, and management and administrative and waste management services	54, 55, 56
Educational services, and health care and social assistance	61, 62
Arts, entertainment, and recreation, and accommodation and food services	71, 72
Other services (except public administration)	81

Notes: This tables shows the mapping between industry titles available in the zip code level data from the ACS on residents' employment by industry, and their asocciated NAICS codes.

Table A3: NYC Cases by Borough

	(1) $\ln(Cases_i)$ NYC	(2) $\ln(Cases_i)$ The Bronx	(3) $\ln(Cases_i)$ Brooklyn	(4) $\ln(Cases_i)$ Manhattan	(5) $\ln(Cases_i)$ Queens	(6) $\ln(Cases_i)$ Staten Island
	Panel A: OLS					
$\%\Delta Trips_i$	0.00963***	-0.00154	0.0172**	-0.00630	-0.00362	-0.00929*
	(0.00345)	(0.00544)	(0.00761)	(0.00954)	(0.00638)	(0.00370)
R-Sq.	0.436	0.635	0.534	0.622	0.262	0.886
Obs.	159	20	34	32	48	10
	Panel B: Reduced form IV					
$ShareTele_i$	-2.761***	-1.789	3.200**	-3.648**	-3.266*	-1.976
	(0.817)	(1.217)	(1.473)	(1.454)	(1.637)	(3.035)
$ShareEss_i$	9.188***	-0.593	9.808***	8.105***	6.720**	-5.566
	(1.103)	(2.227)	(1.912)	(2.476)	(2.557)	(9.172)
R-Sq.	0.675	0.695	0.698	0.725	0.436	0.752
Obs.	159	20	34	32	48	10
	Panel C: IV					
$\%\Delta\widehat{Trips}_i$	0.0605***	-0.000431	0.0435**	0.123	0.0288**	-0.00314
	(0.0126)	(0.00626)	(0.0202)	(0.128)	(0.0124)	(0.00580)
Root MSE	0.443	0.0754	0.252	0.576	0.299	0.0473
Obs.	159	20	34	32	48	10
F-Stat.	35.39	8.172	7.874	2.891	7.908	15.10
	Controls					
$\%AfAm_i$	X	X	X	X	X	X
$ln(Age_i)$	X	X	X	X	X	X
$ln(Inc_i)$	X	X	X	X	X	X

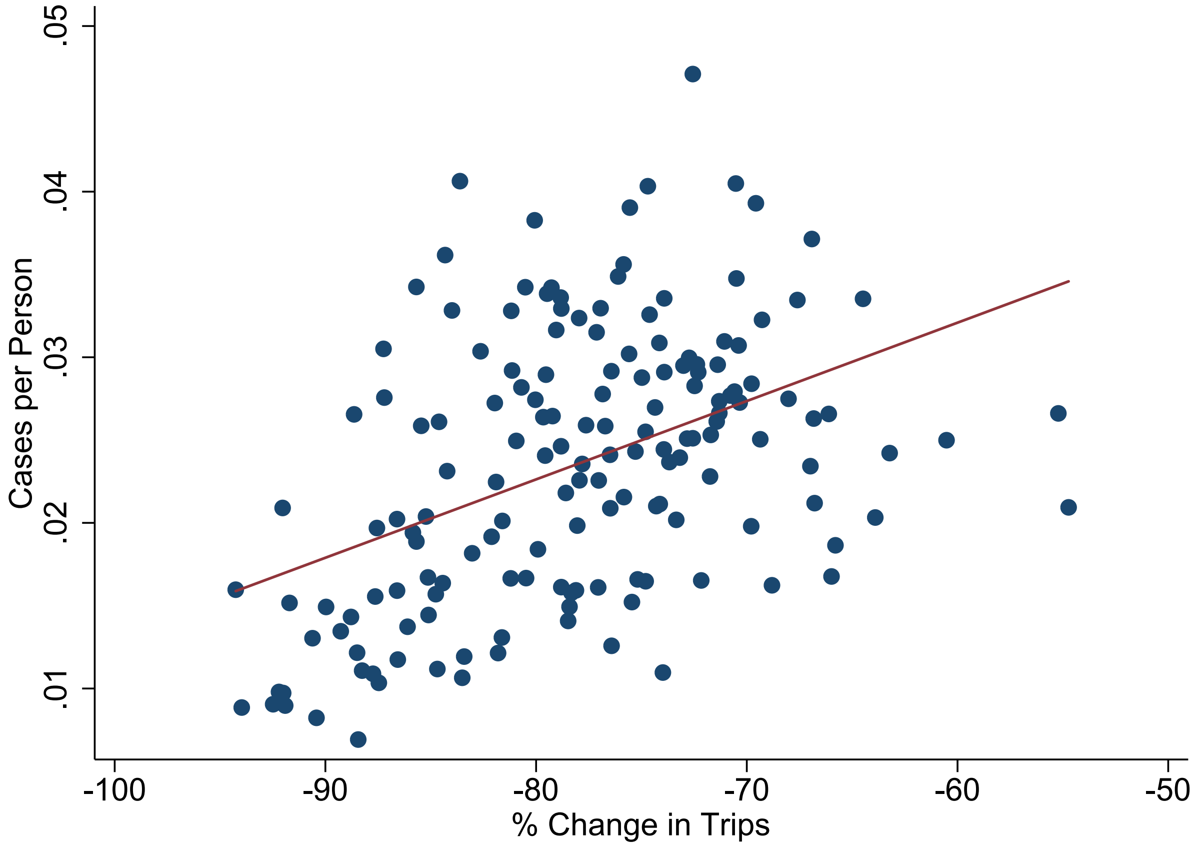
Notes: This table is analogous to Table 2 in the main text, but compares cases across boroughs in NYC instead of different cities. Panels A shows results from Equation (1). Panel B shows the reduced form IV regression results from $ln(TotalCases_i) = \alpha + \beta_1 ShareTele_i + \beta_2 ShareEss_i + \Gamma X_i + \varepsilon_i$. Panel C shows results from Equation (1.2), adding additional demographic controls, X_i: $ln(TotalCases_i) = \alpha + \beta\%\Delta\widehat{Trips}_i + \Gamma X_i + \varepsilon_i$. Robust standard errors in parentheses.
Significance: $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$

Table A4: NYC Deaths by Borough

	(1)	(2)	(3)	(4)	(5)	(6)
	$\ln(Deaths_i)$	$\ln(Deaths_i)$	$\ln(Deaths_i)$	$\ln(Deaths_i)$	$\ln(Deaths_i)$	$\ln(Deaths_i)$
	NYC	The Bronx	Brooklyn	Manhattan	Queens	Staten Island
	Panel A: OLS					
$\%\Delta Trips_i$	-0.00466	-0.0138	0.00189	0.00133	-0.00110	-0.0338**
	(0.00434)	(0.0160)	(0.00892)	(0.0154)	(0.00959)	(0.0103)
R-Sq.	0.476	0.404	0.691	0.638	0.269	0.796
Obs.	159	20	34	32	48	10
	Panel B: Reduced form IV					
$ShareTele_i$	-0.313	-2.949	3.813***	-3.271	-0.626	-1.585
	(1.301)	(3.656)	(1.313)	(2.019)	(3.082)	(14.53)
$ShareEss_i$	5.145**	-3.123	7.134***	15.06***	8.396	-34.22
	(2.108)	(8.025)	(2.363)	(4.992)	(5.496)	(22.10)
R-Sq.	0.507	0.419	0.755	0.745	0.337	0.711
Obs.	159	20	34	32	48	10
	Panel C: IV					
$\%\Delta\widehat{Trips}_i$	0.0269**	-0.00470	0.0287	0.190	0.0261	-0.0426*
	(0.0117)	(0.0216)	(0.0190)	(0.167)	(0.0185)	(0.0220)
Root MSE	0.405	0.217	0.282	0.855	0.386	0.207
Obs.	159	20	34	32	48	10
F-Stat.	40.36	3.075	9.995	3.827	4.667	7.291
	Controls					
$\%AfAm_i$	X	X	X	X	X	X
$ln(Age_i)$	X	X	X	X	X	X
$ln(Inc_i)$	X	X	X	X	X	X

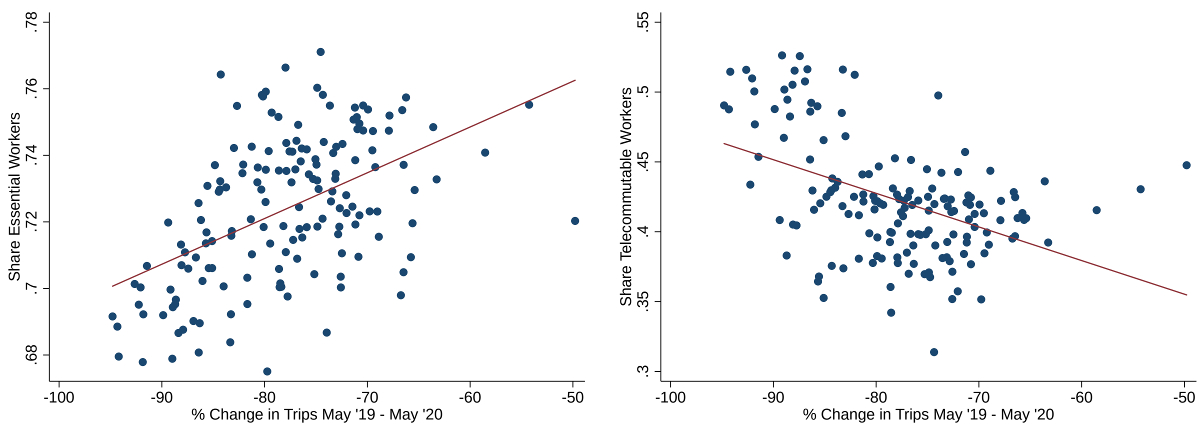
Notes: This table is analogous to Table 2 in the main text, but compares deaths across boroughs in NYC instead of different cities. Panels A shows results from Equation (1). Panel B shows the reduced form IV regression results from $ln(TotalDeaths_i) = \alpha + \beta_1 ShareTele_i + \beta_2 ShareEss_i + \Gamma X_i + \varepsilon_i$. Panel C shows results from Equation (1.2), adding additional demographic controls, X_i: $ln(TotalDeaths_i) = \alpha + \beta\%\Delta\widehat{Trips}_i + \Gamma X_i + \varepsilon_i$. Robust standard errors in parentheses.
Significance: $^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$

Figure A1: Correlation between Travel Change and COVID-19 Cases per Person in NYC



Source: Cases per person from NYC Health Department, available at `https://www1.nyc.gov/site/doh/covid/covid-19-data.page`. % Change in trips from SafeGraph Weekly Patterns Data, using visitors traveling from home. % Change in trips calculated between May 13-19, 2019 and May 4-10, 2020.

Figure A2: A visual first stage
Travel Change and Instruments in NYC



(a) % Δ in Trips vs. $ShareEssential_i$ (b) % Δ in Trips vs. $ShareTelework_i$

Source: % Change in trips from SafeGraph Weekly Patterns Data, using visitors traveling from home. % Change in trips calculated between May 13-19, 2019 and May 4-10, 2020. Share Essential workers calculated from DE and MN 4-digit NAICS essential industries. Share Telework created at the zip level using data from Dingel and Neiman (2020) weighted by local neighborhood employment composition.