Federal Reserve Bank of Chicago

The Relationship between Race, Type of Work, and Covid-19 Infection Rates

R. Jason Faberman and Daniel Hartley

August 2020

WP 2020-18

https://doi.org/10.21033/wp-2020-18

The Relationship between Race, Type of Work, and Covid-19 Infection Rates

August 2020

R. Jason Faberman, Federal Reserve Bank of Chicago
Daniel Hartley, Federal Reserve Bank of Chicago

Abstract
This paper explores the relationship between Covid-19 infection rates, race, and type of work. We focus on three U.S. cities—Chicago, New York, and Philadelphia—allowing us to exploit zip code-level variation in infection rates and testing rates over time, while controlling for a variety of neighborhood demographic characteristics. We find that neighborhoods with higher Black and Hispanic population shares, and neighborhoods with higher shares of workers in high-social contact jobs within essential businesses, had disproportionately higher Covid-19 infection rates, even after applying our testing and demographic controls. These higher rates coincide with citywide peak infection rates, suggesting an amplified response for these groups. Local variation in type of work accounts for relatively little of the variation in infection rates by race. Additional evidence for Arizona, Florida, and Texas also shows amplified infection rates for these groups around statewide peak infection rates, despite their peaks occurring months after the cities in our main sample. Evidence from these states also shows higher infection rates among high-social contact workers in nonessential businesses that coincides with a more aggressive reopening of these businesses.

Keywords: Covid-19 pandemic, infection rates, race, employment, social interaction
JEL Codes: H12, I12, J15, R12

Correspondence: Faberman: jfaberman@frbchi.org, Hartley: Daniel.A.Hartley@chi.frb.org. We thank Michael Fogarty for excellent research assistance. The views expressed here are our own and do not necessarily represent the Federal Reserve Bank of Chicago or the Federal Reserve System.

1. Introduction

By now, it is clear that the Covid-19 pandemic has disproportionately affected minorities in the U.S. Both Black and Hispanic Americans have experienced higher infection rates than White Americans, and in many regions, Black Americans have also experienced higher death rates. It is also clear that the type of work someone does affects their exposure to the disease. Some people work in jobs that can be done from home, while others work in jobs that require being physically close to other people. Some work in industries deemed "essential" during the pandemic. These industries encompass more than health care workers and first responders (for example, grocery store clerks). Others in "nonessential" businesses have seen their industries open up at different times and to differing degrees across the country. Importantly, minorities disproportionally work in jobs that require working in close proximity with others.[1]

In this paper, we examine how race and type of work relate to Covid-19 infection rates. Our goal is to use zip code-level time-series variation in infection rates to quantify their relationship to local demographics—and race in particular—and local employment composition. We focus on three cities where we are able to obtain detailed geographic data over time on the number of Covid-19 cases and Covid-19 testing: Chicago, New York City, and Philadelphia. We match these data to employment and demographic estimates at the zip code level from the American Community Survey. We then estimate the relationship between infection rates—measured as the per capita number of Covid-19 cases—and race and type of work, controlling for zip code-level variation in demographics and weekly testing rates. Controlling for time-variation in testing rates is important due to potential selection effects in who received a test early in the pandemic, when infection rates in all three of our study's cities were high and testing rates were low. Our approach affords us two main benefits. First, it allows us to isolate the

[1] See Mongey, Pilossoph, and Weinberg (2020) for evidence on the demographic characteristics of individuals in jobs that require a high degree of in-person contact and low propensity to work from home.

empirical relationship between infection rates and race and between infection rates and type of work by exploiting zip code-level variation in infections over time. This allows us to control for other confounding factors that may vary by zip code and affect infection rates. Second, it allows us to identify important variation in these relationships over time. The timing of any disproportionate infection rates among minorities or among particular types of workers may coincide with particular aspects of the pandemic, such as periods of citywide high infection rates or the imposition and lifting of stay-at-home orders.

We focus on estimating relative infection rates by race for Blacks and Hispanics, and by type of work for those in *essential* and *nonessential high-social contact* jobs. *Essential* jobs are those within the broad industry categories deemed essential during the stay-at-home order periods across most states and municipalities. *High-social contact* jobs are those that require a high degree of in-person contact and provide a relatively low opportunity to work from home. We find a high degree of residential geographic segregation in the types of jobs people have based on these categorizations in all three cities. Furthermore, the share of a neighborhood's residents employed in high-social contact jobs correlates positively with the neighborhood's infection rates and its share of minority residents across zip codes within each of these cities.

We estimate the relationship between zip code-level infection rates and race and type of work using a panel regression framework. Our analysis shows that, unconditionally, neighborhoods with high shares of workers in high-social contact jobs, and with high shares of Black or Hispanic residents, tended to have disproportionately higher infection rates around the times when citywide infection rates peaked. Neighborhoods with high Hispanic shares and those with high shares of residents in nonessential high-social contact jobs tended to have disproportionately high infection rates for some time after the peak as well. Controlling for neighborhood differences in weekly testing rates and other demographic characteristics (specifically, age, educational attainment, and household composition) accounts for a sizable fraction of these higher rates, but significant differences remain. Moreover, when we jointly estimate the weekly relationships between infection rates and type of work and infection rates and racial composition, we find that neighborhoods with high essential high-social contact worker shares or high Black population shares continue to exhibit higher infection rates around the time when citywide infection rates peak. A 10 percentage point higher share of residents in essential high-social contact jobs is associated with a 4.1 percentage point higher weekly infection rate during April, after applying all controls. Similarly, a 10 percentage point higher share of Black residents is associated with a 0.5 percentage point higher weekly infection rate during April after applying all controls. We also find that neighborhoods with high Hispanic population shares continue to have higher infection rates during and after peak citywide infection rates. A 10 percentage point higher share of Hispanic residents is associated with a 1.3 percentage point higher weekly infection rate during April *and May* after applying all controls. In contrast, a neighborhood's share of nonessential high-social contact workers switches to having *lower* infection rates (relative to the baseline) for much of the sample period once we apply all controls. To put these effects into perspective, weekly infection rates for the pooled sample averaged 0.3 percent in April and just over 0.1 percent in May. Thus, it is not necessarily the case that high-social contact work accounts for higher infection rates among Blacks or Hispanics. If anything, racial composition accounts for the high infection rates we observe among high-social contact workers. Therefore, factors outside of our analysis that are specific to Hispanic neighborhoods must contribute to the disproportionately high rates observed within them. These factors may include language barriers that inhibit the transmission of vital information about the virus, local socioeconomic conditions that affect the risk of infection, or a propensity for activities that lead to higher social contact outside of work, such as greater use of public transit. To a lesser extent, factors specific to Black neighborhoods, independent of their age, educational, or household composition, also contribute to their higher infection rates. We take no stand on what these factors are, but note them as an important avenue for future research.

We supplement our analysis with additional evidence from Texas, Arizona, and Florida. The data from these states are limited in the time-series and geographic variation we can exploit but provide a useful comparison to our main sample because these states experienced a sharp rise in Covid-19 cases well after infection rates in our main sample's cities had abated. These states also took a more aggressive approach to reopening their economies, with most of their businesses operating in at least a limited capacity by early May—the businesses in our sample cities were only partially open by late June. Nevertheless, in the cross-section, we find relationships between (cumulative) infection rates, employment composition, and racial composition that are comparable to the relationships we document using our main sample. Within Texas, where we have a time series of Covid-19 case and testing data for each county, we find disproportionately high infection rates in counties with high minority population shares or high shares of nonessential high-social contact workers. As with our main analysis, we continue to find significantly higher infection rates for these groups even after applying all of our controls. While peak infection rates in Texas occur in June through early July, we find similar quantitative effects by race. After applying controls, we find that a 10 percentage point increase in a county's Black population share is associated with a 0.8 percentage point higher weekly infection rate, and a 10 percentage point increase in its Hispanic population share is associated with a 1.5 percentage point weekly higher infection rate during the period of rising statewide infection rates. We also find disproportionately higher infection rates for nonessential high-social contact workers during this period. In Texas, peak infection rates coincided with the aggressive reopening of nonessential businesses, which may explain the differences between these findings and those for our main sample. A 10 percentage point increase in the share of nonessential high-social contact workers is associated with a 4.6 percentage point higher weekly infection rate. Overall, this supplemental evidence reinforces the findings from our main analysis. Specifically, that periods of high infection rates amplify the risk faced by those most exposed to the virus, and that demographics and type of work alone cannot explain the high infection rates Blacks and Hispanics have experienced during these peak periods.

Our study is one of many recent studies to examine the relationship between economic activity and Covid-19 infections. Several studies have examined the correlations between Covid-19 outcomes (either infection rates or mortality rates) and local socioeconomic and demographic characteristics. These studies predominantly focus on county-level relationships between cumulative Covid-19 outcomes and local characteristics in the cross-section rather than the within-area time-series variation we exploit here.[2] Nevertheless, these studies cover the entire nation and consistently find racial disparities in Covid-19 outcomes. Benitez, Courtemanche, and Yelowitz (2020) relate zip code level variation in Covid-19 cases to a variety of local demographic and socioeconomic characteristics. Like our study, they find that these characteristics can only partially account for racial disparities in Covid-19 case rates, but they only examine these relationships in the cross-section. Papageorge et al. (2020) show that socioeconomic conditions are strongly tied to one's propensity to engage in social distancing and other protective behavior—those in worse-off socioeconomic conditions are *less* likely to engage in social distancing behavior. Glaeser, Gorback, and Redding (2020) have a study most similar to ours. They examine five U.S. cities and exploit zip code-level variation over time to estimate the relationship between mobility (based on cell phone data) and Covid-19 infection rates, using local employment composition as an instrument. They find a strong relationship between mobility and infection rates, and similar to our study, find that the relationship is strongest during peak citywide infection rates.

The next section describes our data and measurement, and presents summary evidence for how Covid-19 cases have evolved in our three sample cities. Section 3 presents our main evidence on the

[2] These studies include Brown and Ravallion (2020), Chen and Kreiger (2020), Knittel and Ozaltun (2020), and McLaren (2020).

relationship between infection rates, race, and type of work, as well as the supplementary evidence from Arizona, Florida, and Texas. Section 4 concludes.

2. Data and Measurement

We match weekly data on Covid-19 cases and testing to aggregated employment and demographic data from the 2014-18 American Community Survey (ACS) at the zip code level for three cities: Chicago, New York City, and Philadelphia. All three cities have publicly-available, frequent updates to their Covid-19 data at the zip code-level. The ACS provides aggregate estimates of employment in broad industry and occupation categories and population totals by race, education, and age, and other demographic characteristics.

Our analysis focuses on the relationship between the racial and employment make-up of each zip code and Covid-19 infection rates over time, controlling for local testing rates and other demographic characteristics. For employment, we focus on a distinction between jobs that require high social contact with other people and those that do not. We do this since increased contact with others increases the chances one will contract Covid-19. We also distinguish between jobs that states classified as essential services and jobs states classified as nonessential. Those working in essential industries were exempt from stay-at-home orders throughout the pandemic and often were required to work, while those working in nonessential industries incrementally returned to work as each city gradually reopened.

We classify jobs as either high or low social proximity based on the social proximity index derived by Leiboivci, Santacreu, and Famiglietti (2020).[3] Their index uses job task information from the O*NET database of occupations to create an index of the degree of social contact individuals typically make while on the job. We then use their proximity index at the two-digit Standard Occupational Classification

[3] We thank Fernando Leibovici for generously providing us with their proximity estimates.

(SOC) level and interact it with estimates of the fraction of each occupation that can plausibly work from home, as derived by Dingel and Neiman (2020). They also use job task information from O*NET to derive their estimates.[4] This gives us an *effective proximity index* for each occupation. Letting PI_j denote the proximity index for occupation j from Leibovici et al. and WFH_j denote the work-from-home share for occupation j from Dingel and Neiman, our effective proximity index equals $PI_j(1 - WFH_j)$. The effective proximity index captures the fact that many individuals who have been able to work from home have done so during the crisis, mitigating their social contact on the job.

We then classify broader one-digit occupations as either high-social contact or low-social contact based on the effective proximity index estimates of their two-digit occupations. We must do this because the employment data in the ACS are only available for broad industry and occupation categories at the zip code level. As it turns out, nearly all broad occupation categories contain two-digit occupations that are all high-social contact or all low-social contact, as Table 1 shows. There are a few notable exceptions. Healthcare practitioners are a high-social contact occupation, but make up a small fraction of the Management, Business, Science, and Arts occupation category (which is otherwise a low-social contact category), and are a minority of the group's employment even within the Education and Health industry sector. Thus, we count the Management, Business, Science, and Arts occupation category as low-social contact across all sectors. The farming, fishing, and forestry occupations are relatively low-social contact, but the remainder of the Natural Resources, Construction, and Maintenance occupation category is high-social contact. Again, this occupation makes up a minority of the broader category's employment, so we classify the group as high-social contact. The exception is within the Mining and Logging industry sector, where farming, fishing, and forestry occupations make up

[4] The estimates of social proximity from Leiboivci, Santacreu, and Famiglietti (2020) are very similar to those generated by Mongey, Pilossoph, and Weinberg (2020). Several studies also find similar work-from-home estimates to Dingel and Neiman (2020). These include Aaronson, Burkhardt, and Faberman (2020), Bartik et al. (2020), and Brynjolfsson et al. (2020).

the majority of the group's employment. For this sector, we classify the Natural Resources, Construction, and Maintenance occupation group as low-social contact. In practice, this is not a relevant sector for our analysis since we focus on large urban areas. Table 1 shows that, among the remaining occupation categories, service occupations (which include healthcare support, protective services, food and serving related jobs, maintenance jobs, and personal service jobs) and production and transportation-related jobs are the other high-social contact occupation categories in our analysis.

We classify jobs as essential or nonessential based on the share of employment in each broad industry sector identified as essential by Aaronson, Burkhardt, and Faberman (2020). They use a detailed listing from Massachusetts to impute an essential-worker employment share for each three-digit NAICS industry. We calculate the employment-weighted average of their estimates for each broad industry sector observed in the ACS, and report these estimates in Table 2.[5] We establish as our cutoff that each broad industry sector has to have at least 80 percent of its employment deemed essential to count as an essential sector in our study. The six sectors that meet this criterion are 1) Construction; 2) Manufacturing; 3) Transportation, Warehousing, and Utilities; 4) Finance, Insurance, and Real Estate; 5) Education and Health; and 6) Public Administration. For reference, Table 2 also reports the (employment-weighted) average effective proximity index for each industry sector. There is little relation between the average index value and whether or not an industry sector is essential, underscoring our need to account for both industry and occupation variation in employment across zip codes. In our analysis, we focus on the zip code-level employment shares of three groups of workers:

[5] We use employment estimates from the February 2020 Current Employment Statistics survey to generate the sectoral-level essential worker shares and employment estimates from the 2019 Occupational Employment Statistics to generate the sectoral-level effective proximity index estimates. We also note that, although there were variations in what counted as essential businesses across states, most of these differences occur well below the one-digit industry categorization we use in our analysis.

essential workers in high-social contact occupations, nonessential workers in high-social contact occupations, and low-social contact workers, regardless of essentiality.

We obtain demographic data for each zip code from publicly-available statistics on zip code-level population counts from the American Community Survey (ACS). From these counts, we generate the population shares by race, age, educational attainment, and household composition. The data are generated from pooled ACS surveys between 2014 and 2018. We also generate employment shares by broad industry × occupation sector for each zip code from the ACS data. We use these shares to calculate the fraction of zip-code level employment in essential vs. nonessential and high-social contact vs. low-social contact jobs. Note that these employment shares are based on workers' location of residence rather than location of work.

Finally, we obtain data on Covid-19 infection and testing rates from the Public Health department websites of the three cities in our study: Chicago, New York City, and Philadelphia.[6] We chose these cities primarily because they have the most comprehensive, publicly available data on Covid-19 infections and testing at the zip code level at a high frequency. For each city, the data include the total number of Covid-19 tests in a given week, the number of cumulative tests through the end of that week, and the total number of positive tests (infections) in a given week and cumulatively. Our main variable of interest is the Covid-19 infection rate, which we measure as the number of positive test cases reported in a given week per 100,000 population. Since positive test results can arise because of increased testing over time, we control for testing rates throughout much of our analysis and measure these rates as the number of tests conducted in a given week per 100,000 population. Because of differences in data availability, our sample periods vary by city. At the zip code-level, our data for

[6] Our data for Chicago are from the City of Chicago Department of Health and Human Services and cover the city limits within Cook County. Our data for New York City come from the New York City Department of Health and Mental Hygiene and cover the five boroughs (Manhattan, Bronx, Brooklyn, Queens, and Staten Island). Our data for Philadelphia come from the City of Philadelphia Department of Public Health and cover the city of Philadelphia.

Chicago begin the week ending March 21, 2020, our data for New York City begin the week ending April 4, 2020, and our data for Philadelphia begin the week ending May 2, 2020. At the city level, we have aggregate data that go back to at least March 21 for all three cities.

Figure 1 shows the patterns of weekly infection rates, testing rates, and share of tests that were positive for Covid-19 for each city in the sample. The three cities differ in the magnitudes of each variable, as well as the timing of their peaks, but all three cities share the same qualitative patterns. Specifically, all three cities experience an increase in their Covid-19 infection rates that peaks sometime in April or early May and steadily declines thereafter. New York City has the highest peak infection rate, which occurs in early April. It also generally has higher infection rates until about June, when its rate continues to fall while the rates of Chicago and Philadelphia level off. Chicago and Philadelphia have peak infection rates that are similar in magnitude, though Chicago reaches its peak in early May while Philadelphia reaches its peak in mid-April. In all three cities, testing rates steadily rise throughout the sample period. In Chicago, testing rates peak in late May then level off. In New York, testing rates rise throughout, with a notable ramping up in mid-May. In Philadelphia, testing rates rise throughout the sample period at a rate comparable to Chicago's testing rate. Finally, weekly positive Covid-19 test rates rise then fall relatively sharply in each city. New York City exhibits the highest positive test rates early on, peaking at 65 percent in early April. As testing increases in New York City, however, the fraction that are positive falls relatively quickly and is below the positive test rates of the other two cities by the end of the sample period. Weekly positive test rates follow a similar pattern in Chicago and Philadelphia, and peak in mid-April in both cities. The peak for Philadelphia is somewhat higher (39 percent) than the peak for Chicago (34 percent).

3. Evidence

3.1. Motivating Evidence

We begin our analysis with some motivating evidence for jointly studying race and employment as it relates to the Covid-19 pandemic. First, we quantify the differences in infection rates by race. Table 3 reports cumulative infection rates of Blacks and Hispanics relative to Whites. We report the cumulative infection rates (measured as total cases per capita) through the week of July 11. We do this because each city differs in the timing of available Covid-19 data by race. Furthermore, we normalize the cumulative infection rates within each city by the infection rate for Whites. We do this because of the wide variation in these rates, across all races, across cities. Table 3 shows that both Blacks and Hispanics exhibit relatively higher infection rates than Whites across all three cities. The Black infection rate is 1.5 to 2.4 times higher than the White infection rate, and the Hispanic infection rate is 1.2 to 3.7 times higher than the White infection rate. The disparities are greatest in Chicago and smallest in New York City.

Next, we highlight how race varies with type of work. To do so, we pool all workers from the 2019 monthly Outgoing Rotation Groups of the Current Population Survey and generate estimates for the entire United States. We then estimate the share of workers who are Black or Hispanic within our four broad employment categories: 1) essential high-social contact jobs, 2) nonessential high-social contact jobs, 3) essential low-social contact jobs, and 4) nonessential low-social contact jobs. Table 4 presents the estimates. Black workers make up 12.2 percent of total employment but 16.1 percent of essential, high-social contact jobs. In contrast, they only make up 9.8 percent of nonessential, low-social contact jobs. Hispanic workers make up a disproportionate share of workers in high-social contact jobs, regardless of essentiality. They account for 17.6 percent of total employment, but 24.6 percent of essential high-social contact jobs and 27.2 percent of nonessential high-social contact jobs. Hispanics are least represented in essential, low-social contact jobs, making up only 11.7 percent of its employment. For reference, essential jobs make up just over 57 percent of total employment, while high-social contact jobs make up just over 37 percent of total employment.

Figure 2 shows the potential importance of type of work for the geographic distribution of infection rates. For each city, it maps the share of each zip code's employment in high-social contact jobs (in both essential and nonessential businesses), based on the jobs of that zip code's residents. The key implication from the figure is the stark geographic dispersion of workers by their type of job within each city. Those who live in the central business districts of each city are disproportionately in low-social contact jobs. These include the Downtown and Loop areas in Chicago, Manhattan and parts of Brooklyn in New York City, and Center City in Philadelphia. In contrast, those who live further from the downtown areas are disproportionately in high-social contact jobs. If these jobs are located in the central business districts at least as much as they are located throughout the remainder of each city, it would suggest that workers who reside outside of the central business district are more likely to use public transit to get to work, and therefore have even higher rates of contact with others than even their job duties imply.

Table 5 shows that in the cross-section zip codes with the highest shares of high-social contact workers also have the highest shares of minorities *and* the highest Covid-19 infection rates. The table presents the univariate (population-weighted) correlations between race, type of work, and Covid-19 infection rates across zip codes within our three sample cities and within the zip codes pooled across the three cities. The top panel reports the correlations between racial shares (percent of each zip code that is Black or Hispanic) and employment shares by type of work (percent of each zip code's residents employed in essential or nonessential high-social contact jobs). The top panel shows that the national patterns reported in Table 3 are also present geographically in our sample cities. Zip codes with high Black population shares also have high employment shares in essential, high-social contact work. Zip codes with high Hispanic population shares also have high employment shares in both essential and nonessential high-social contact work. The bottom panel of Table 5 shows the correlations of the racial and employment shares with Covid-19 infection rates across zip codes. We measure the infection rate as the cumulative number of positive cases per 100,000 population through July 11, 2020 (the end of our sample period). The bottom panel shows a strong positive correlation between employment shares in essential, high-social contact jobs and infection rates across all three cities. It also shows a strong correlation between nonessential, high-social contact shares and infection rates in Chicago and New York City, with a weaker but still positive correlation in Philadelphia. The share of each zip code that is Black is weakly positively correlated with infection rates overall, though there is wide heterogeneity across cities—in Philadelphia, the correlation is 0.45 while in Chicago the two are essentially uncorrelated. The correlation of the Covid-19 infection rate with share of each zip code that is Hispanic also has considerable heterogeneity across the three cities. The correlation is strong and positive in Chicago and New York City (0.74 and 0.46, respectively), but the two are essentially uncorrelated in Philadelphia.

Thus, in the cross-section of zip codes both within and across our sample cities, there are strong, positive joint relationships between racial composition, the composition of the type of work individuals do, and infection rates. Minorities disproportionately work in jobs that require high-social contact with others and live in neighborhoods where workers in these types of jobs are overrepresented. With few exceptions, these are also the neighborhoods with the highest Covid-19 infection rates.

3.2. Panel Data Analysis

We now move on to our main analysis of the relationships of Covid-19 infection rates with type of work and race. We exploit zip code-level time-series variation in our data to estimate these relationships. Note that our estimates do not reflect causal effects on infection rates. Instead, they are the conditional correlations with infection rates controlling for all other factors in our regression model. We estimate our main specification pooling the data for all three cities together, though we also present results for each city separately below. We evaluate the relationship between infection rates and type of work over time, and between infection rates and racial composition over time, controlling for weekly variation in zip-code level testing rates and other demographic characteristics of the zip code. Our regression specification is

$$C_{ijt} = \alpha_{jt} + \eta T_{ijt} + \beta_t^E EHC_{ij} + \beta_t^N NHC_{ij} + \gamma_t^B B_{ij} + \gamma_t^H H_{ij} + X_{ij}\delta + \varepsilon_{ijt}, \quad (1)$$

where the weekly infection rate for zip code i in city j in week t is C_{ijt} and we measure it as the number of positive cases per 100,000 population during that week. We include a set of city-specific week effects, α_{jt} and the weekly Covid-19 testing rate within each zip code, T_{ijt}, measured as the number of tests administered per 100,000 population during that week. The first set of coefficients of interest are β_t^E and β_t^N, which estimate the week-specific relationship of the share of workers in essential high-social contact jobs, EHC_{ij}, and the share of workers in nonessential high-social contact jobs, NHC_{ij}, in zip code i in city j, respectively, by interacting each share with week fixed effects. Implicitly, the model estimates these coefficients relative to the weekly infection rates for all low-social contact jobs. The second set of coefficients of interest are γ_t^B and γ_t^H, which estimate the week-specific relationship of the share of each zip code that is Black, B_{ij}, or Hispanic, H_{ij}, respectively, by interacting each share with week fixed effects. Finally, we include additional zip code-level demographic controls, X_{ij}. These include the share of each zip code within three age groups, two educational attainment groups, and three groups for the number of workers in each household.[7] We cluster standard errors by zip code and weight the regression by zip code population.

[7] Specifically, we include controls for the share of the population age 18 to 39, age 40 to 64, and age 65 or more, the share of the population with a high school degree or less, or with some college, and the share of the population with one worker, two workers, or three or more workers in the household. We also experimented with

By estimating time-varying relationships between infection rates, employment shares, and racial shares, we can examine how these relationships varied in conjunction with changes in the severity of Covid-19 infections and the related imposition and relaxation of stay-at-home orders within these cities during our sample period. Controlling for weekly testing rates allows us to account for potential selection of identified positive cases—access to testing in most areas has been uneven and nonrandom. The issue was most pronounced early in the pandemic, when testing rates were low across the nation. Finally, our additional demographic controls allow us to account for other factors that we suspect may affect infection rates. National evidence on Covid-19 cases shows that younger individuals account for a disproportionate share of infections, despite older individuals accounting for a disproportionate share of hospitalizations and deaths. National evidence on Covid-19 cases also shows that infection rates are higher among the less educated. Finally, there is a concern that community spread of the virus can be greater in areas with larger households. To the extent that the virus could spread to household members by those who contract it through their jobs, our controls for the number of workers per household will capture this potential for community spread.

We present the estimates from equation (1) in Figures 3 and 4. Figure 3 presents the coefficient estimates for β_t^E and β_t^N from equation (1) for four different regression specifications. The first, unconditional, specification includes the weekly interactions with the employment shares, EHC_{ij} and NHC_{ij}, and controls for week × city effects, α_{jt}, but nothing else. The second specification additionally controls for weekly testing rates, T_{ijt}. The third specification additionally controls for the additional demographic variables (age composition, educational composition, and household composition). The fourth specification is the full model specified in equation (1) and thus additionally includes the weekly

using the number of household members rather than the number of workers, but the latter had a stronger relationship to infection rates in all of our regression estimates.

interactions with racial composition. The top panel of Figure 3 reports the estimates of β_t^E for each specification, while the bottom panel reports the estimates of β_t^N for each specification.

The panels of Figure 3 show that, unconditionally, zip codes with higher shares of both essential and nonessential high-social contact workers had disproportionately higher Covid-19 infection rates, relative to areas with higher shares of low-social contact workers. These higher infection rates differ in their magnitude and in their timing. Areas with higher shares of essential high-social contact workers had weekly infection rates that were 1,080 cases per 100,000 population higher than the baseline rate at their peak, which occurred in early April. Specifically, a 10 percentage point increase in the share of a zip code's employment in essential high-social contact work was associated with up to a 10.8 percentage point higher infection rate in that zip code. The fact that the relative infection rates are largest during the period when infection rates were surging for all areas in our sample suggests that this surge was *amplified* in areas with higher shares of essential high-social contact workers. Following the surge periods, relatively higher infection rates in these areas dissipate and are essentially the same as the baseline infection rates by the beginning of June. For zip codes with a higher share of nonessential high-social contact workers, a 10 percentage point increase in the share of these workers was associated with up to a 5.7 percentage point higher infection rate. This peak occurred later, in early May, and followed several weeks of essentially zero difference with the baseline infection rate. It remained somewhat elevated—at around a 1 percentage point higher weekly infection rate—through the beginning of July. These results are consistent with what one might expect given the nature of the stay-at-home orders and definition of essential work during these times. The strictest stay-at-home orders remained in place through early-to-mid June in our three sample cities. During these orders, essential workers were the only ones at work. Furthermore, many, such as first-responders and those in health care, were dealing with the Covid-19 pandemic directly. Thus, one might expect an amplification of their infection rates relative to others during this period. As infection rates fell and our sample cities lifted their most restrictive stay-at-home orders, infection rates among essential high-social contact workers fell while infection rates among nonessential high-social contact workers rose (relative to the baseline rates). The timing of the peak in infection rates among zip codes with high shares of nonessential high-social contact workers is difficult to explain, since it occurs about one month before most stay-at-home orders were lifted, and consequently when many of these workers returned to work. Keep in mind, though, that these are our unconditional estimates and do not control for the other factors in our model that might affect infection rates independently of the stay-at-home orders.

Controlling for changes in zip code-level testing rates has a notable impact on our estimates for the relationship between infection rates and the local share of essential high-social contact workers. In the early period, when infection rates were surging and stay-at-home orders were in place, it reduces our estimated coefficients by about one-quarter, but actually increases our coefficient estimates thereafter. In fact, controlling for testing rates suggests a steadily increasing relationship between a higher share of essential high-social contact workers and weekly infection rates from the beginning of June onward. Unconditionally, a 10 percentage point increase in a zip code's share of essential high-social contact workers was associated with a 0.3 percentage point higher infection rate from June forward, but after controlling for local testing rates, it is associated with a 2.9 percentage point higher infection rate. Keep in mind that infection rates are falling in our sample cities during this period, so the coefficients suggest that infection rates fell more slowly in neighborhoods with more essential high-social contact workers. Controlling for testing rates has a small effect on the relationship between infection rates and a zip code's share of nonessential high-social contact workers, reducing our coefficient estimates an average of 17 percent across all weeks. The time-series pattern remains generally unchanged, save for the fact that higher shares of nonessential high-social contact workers are no longer associated with higher infection rates after the end of May.

Additionally controlling for age, education, and household composition reduces the magnitudes of the relationships of essential and nonessential high-social contact worker shares to infection rates somewhat, but their time-series patterns remain essentially unchanged. It reduces our coefficient estimates on essential high-social contact worker shares by about 30 percent, on average, and it reduces our coefficient shares on nonessential high-social contact shares so that they have close to a zero average relationship.[8] As Figure 3 shows, though, the zero-average relationship masks estimates of relatively higher infection rates in the first half of our sample period and relatively lower infection rates in the second half of our sample period. Finally, additionally controlling for the time-varying relationships between a zip code's racial composition and infection rates has little effect on our coefficient estimates for essential high-social contact worker shares but a sizable effect on our coefficient estimates for nonessential high-social contact worker shares. In fact, controlling for the time-varying effects of racial composition drives the relationship between infection rates and the share of nonessential high-social contact workers to be slightly negative over the sample period, relative to the baseline infection rates. After adding all controls, a 10 percentage point increase in the share of nonessential high-social contact workers is associated with a 1.5 percentage point *lower* infection rate, on average, with the lowest relative rates in June. We return to this point in our discussion of the racial composition results below.[9]

Figure 4 presents the coefficient estimates for γ_t^B and γ_t^H from (1) for four different regression specifications, specified analogously to those in Figure 3. The first, unconditional, specification includes

[8] In the appendix, we report the coefficient estimates associated with our additional controls. We find a strong positive relationship between testing rates and infection rates. We find little relation between infection rates and a zip code's age distribution. We find that a higher share of those with a high school degree or less is related to higher infection rates, as are higher shares of households with at least one member employed.

[9] In the appendix, we report the estimates from Figures 3 and 4 that include our additional controls but not the joint estimates of weekly type of work and race relationships and the estimates from our full regression model with 95 percent confidence intervals included. The appendix figures show that, in general, the relatively high infection rates for zip codes with relatively high shares of essential high-social contact workers, Blacks, and Hispanics highlighted in the main text are all significantly different from zero.

the weekly interactions with the race shares, B_{ij} and H_{ij}, and controls for week × city effects, α_{jt}, but nothing else. The second specification additionally controls for weekly testing rates, T_{ijt}. The third specification additionally controls for the additional demographic variables (age composition, educational composition, and household composition). The fourth specification is the full model in (1) and thus additionally includes the weekly interactions with the employment shares—consequently, the estimates from the full specification in Figures 3 and 4 come from the same regression. The top panel of Figure 4 reports the estimates of γ_t^B for each specification, while the bottom panel reports the estimates of γ_t^H for each specification.

The panels of Figure 4 show that there are disproportionately high infection rates among zip codes with high Black or Hispanic population shares. As with the employment shares, there are distinct patterns in the timing of when areas with high Black or Hispanic population shares experience relatively high infection rates. Unconditionally, zip codes with a higher Black population share had disproportionately higher infection rates throughout April until early May. These rates peak in late April when a 10 percentage point higher Black population share is associated with a 2.0 percentage point higher infection rate, then return to have essentially no difference with the baseline infection rates from June forward. Zip codes with a higher Hispanic population share follow a similar qualitative pattern, but with notably larger and more persistent effects. At its peak in late April, a 10 percentage point higher share of a zip code's Hispanic population is associated with a 3.7 percentage point higher infection rate. A higher Hispanic population share remains associated with higher infection rates throughout April and May, with a 10 percentage point higher share associated with a 1 to 3 percentage point higher infection rate throughout this period. The effects of a higher Hispanic population share fall further following the end of May, but remain positive through the end of the sample period. Thus, at least unconditionally, there appears to be amplification in infection rates by race as well. Neighborhoods with high Black and Hispanic population shares experienced disproportionately high infection rates as overall infection rates were surging and stay-at-home orders were in place. High infection rates among predominantly Hispanic neighborhoods persisted further into the times where stay-at-home orders were partially lifted.

Controlling for changes in zip code-level testing rates leads only to a small reduction in the estimated positive relationships between infection rates and racial composition. The largest reductions are during the early part of our sample period, when it reduces our coefficient estimates by about 15 percent for both Black and Hispanic population shares. Similar to its effect on our results for essential high-social contact workers, controlling for testing rates implies a small but gradual increase in relative infection rates among zip codes with higher Black population shares from the early June forward.

Additionally controlling for age, education, and household composition leads to somewhat larger reductions in our estimated coefficients. As with the employment shares, however, their time-series behavior is qualitatively the same. Notably, though, controlling for the additional demographics reduces the relationship of infection rates to the Black population share to about zero by early May and reduces their relationship to the Hispanic population share to about zero by early June. Also notable is that, even after controlling for these demographic characteristics, zip codes with high Hispanic population shares continue to have disproportionately high infection rates in the early months of our sample period. There has been speculation among policymakers and health professionals that the greater propensity of larger and multigenerational families among Hispanic households may be a key contributor to their higher infection rates. Our estimates suggest that household composition, along with differences in the age and educational attainment composition of their neighborhoods, can only account for about 40 percent of the higher infection rates we observe in zip codes with high Hispanic population shares during April and May.[10]

---

[10] Specifically, when we control for testing rates, a 10 percentage point increase in a zip code's Hispanic population share is associated with a 2.1 percentage point higher infection rate, on average, during April and May. When we

Finally, additionally controlling for the time-varying relationships between employment shares and infection rates has an impact on the coefficients on Black population shares in the early weeks of our sample, reducing the coefficient estimates by over one-half, but almost no impact on their estimated effects afterward. Note that the estimates associated with these controls are the same as those reported for the full specification in Figure 3. In contrast, controlling for the time-varying relationships between employment shares and infection rates has almost no effect on the higher infection rates associated with Hispanic population shares. This is notable because our motivation for this study is to examine the extent that the type of work minorities do affects their infection rates. The estimates suggest that the propensity of Blacks to work in essential high-social contact jobs accounts for about one-third of the unconditional relationship between higher infection rates and the Black population share during the early weeks of our sample. In contrast, the propensity of Hispanics to work in essential and nonessential high-social contact jobs accounts for almost none of the higher infection rates we observe in neighborhoods where the Hispanic population share is high, conditionally or unconditionally, throughout our sample period.

Returning to our estimates in Figure 3, controlling for time-varying relationships of racial composition and local infection rates has a sizable effect on the estimated impacts of type of work, not the other way around. Controlling for race essentially wipes out the positive association between the share of workers in nonessential high-social contact jobs and weekly infection rates, switching the relationship from being somewhat positive, particularly between mid-April and the end of May, to being somewhat negative throughout our sample period. Therefore, our estimates suggest that the main reason that we find higher infection rates in areas with a higher share of nonessential high-social contact workers is because those jobs are disproportionately filled by Hispanic workers.

additionally control for age, education, and household composition, a 10 percentage point increase in this share is associated with a 1.3 percentage point higher infection rate during these months.

Summarizing our main results: unconditionally, we find that neighborhoods with high shares of high-social contact workers and high shares of Black or Hispanic residents had higher infection rates across our three sample cities, particularly in April and much of May. Controlling for local variation in testing rates and demographics accounts for about 55 percent of the disproportionately high infection rates in neighborhoods with higher shares of essential high-social contact workers and over 70 percent of the disproportionately high infection rates in neighborhoods with high Black population shares during this period. It also accounts for roughly half of the disproportionately high infection rates in neighborhoods with high Hispanic population shares and more than all of the disproportionately high infection rates in neighborhoods with high shares of nonessential high-social contact workers during this period. For the most part, there is little relation between these neighborhood characteristics and infection rates after the end of May, when infection rates fell throughout all three cities in our sample and stay-at-home orders were gradually lifted. Importantly, even after adding controls, we continue to find significantly higher infection rates in neighborhoods with high Black population shares and high shares of essential high-social contact workers in April, and significantly higher infection rates in neighborhoods with high Hispanic population shares in April and May. Just under half of the higher infection rates within predominantly Hispanic neighborhoods are due to differences in testing rates, household composition, the types of jobs held by their residents, or other demographic characteristics. Other factors must account for the remainder. These may include language barriers that may inhibit the transmission of vital information on protecting oneself from Covid-19, socioeconomic conditions outside of our controls that put Hispanics at a relatively greater risk of infection, or a propensity for higher social contact outside of work (for example, because of a higher likelihood of using public transit). To a lesser extent, similar factors likely account for higher infection rates among neighborhoods with a high Black population share during the peak of the pandemic in these cities.

3.3. Supplemental Evidence

While our three cities allow for a robust estimation of the relationships between race, employment, and infection rates, they are not representative of the experience of the Covid-19 pandemic across the nation. From March through May, Covid-19 cases surged in New York City and throughout much of the Northeastern U.S., as well as larger cities such as Chicago and Los Angeles. The surge in New York City was particularly acute, and during the first two months of the pandemic, the city accounted for about one-third of all Covid-19 cases in the U.S. Since June, however, all three cities in our sample have seen their infection rates fall dramatically, while other parts of the U.S., particularly in the South and West, have seen their infection rates surge.

In this subsection, we attempt to evaluate how representative our results are for other areas. First, we re-estimate our regression model specified in equation (1) separately for each sample city to examine the degree to which our pooled results obscure important heterogeneity across the cities. Next, we appeal to available data for three large states where infection rates surged starting in June. We are limited in the type of data available (otherwise we would have included them in our main analysis), so we present evidence in as consistent a manner that each state's data allow to see how it compares to our main estimates. For Texas, we have a time series on *county*-level Covid-19 infection and testing data from late April forward, so we use these data matched to the ACS data to estimate equation (1) at the county level. For Arizona and Florida, we only have a cross section of Covid-19 cases by zip code, so we match these to the ACS data and estimate cross-sectional correlations comparable to those we report in Table 5.

Separate Estimates by City. One might worry that our pooled estimates mask important heterogeneity in our estimated relationships across cities. After all, Figure 1 shows that each city experienced notably different patterns in the rise and fall in their infection rates. We therefore estimate our regression model in equation (1) separately for each city. The specifications are the same as before except that we control for week effects, α_t, rather than city × week effects, α_{jt}, within each city-specific regression.

Our results are shown in Figures 5 and 6. We report the estimates for the full specification in equation (1) for each city in an analogous manner to how we report them in Figures 3 and 4, so for each city, the coefficients in both figures come from the same regression. The two panels of Figure 5 report the city-specific estimates for β_t^E and β_t^N and the two panels of Figure 6 report the city-specific estimates for γ_t^B and γ_t^H. The figures show that there is notable heterogeneity in the timing and patterns of the relationships of employment and racial composition with infection rates across the cities, but there are notable similarities as well. Figure 5 shows that once we control for all factors in our regression specification, the bulk of the disproportionately high infection rates among zip codes with high shares of essential high-social contact workers occurs in New York City. We note, though, that our time series for Philadelphia begins just after it experiences its peak infection rates. A gradual rise in infection rates within these zip codes occurs in both New York City and Chicago from the end of May forward (again, implying that infection rates fell more slowly in these areas than the city as a whole, since overall infection rates declined during this period). Once we control for all other factors, both New York City and Philadelphia show either no different or slightly lower infection rates for zip codes with higher shares of nonessential high-social contact workers. Chicago exhibits somewhat higher infection rates in April and somewhat lower infection rates in June for these zip codes. Otherwise, the patterns across the three cities are similar.[11]

Figure 6 shows that all three cities have very similar patterns for relative infection rates for zip codes with higher Black population shares. Of the three, Chicago shows somewhat higher infection rates

[11] For reference, infection rates peaked at 250 per 100,000 population for Chicago the week ending May 2, at 450 per 100,000 population for New York City the week ending April 4, and at 230 per 100,000 population for Philadelphia the week ending April 11 (see Figure 1).

within these zip codes in April and early May. All three cities show a peak in disproportionately high infection rates among zip codes with high Hispanic population shares, and there is some persistence in these peaks in all three cities, though the timing differs. In New York City, these zip codes have higher rates in early April, concurrent with peak infection rates for the city overall. The coefficient estimates fall over time, but remain elevated through late May. Chicago has a larger peak in relative infection rates. It occurs later, but is also concurrent with Chicago's overall peak infection rate in early May and declines gradually through the end of May. Philadelphia is the only city whose relatively high infection rates within these zip codes occurs after its peak infection rate. Infection rates within these zip codes in Philadelphia are somewhat higher (relative to the baseline) in May and most of June, but are then comparable to baseline infection rates thereafter.

Estimates for Texas Counties. Our next exercise examines how our results look for a region with an entirely different experience with the pandemic. Infection rates in Texas rose gradually through March but remained relatively low through April and May. Texas instituted a statewide stay-at-home order for the month of April, but lifted the order thereafter and returned to close to a complete reopening in many parts of the state at the beginning of May. By early June, infection rates across the state began to rise dramatically, and continued to increase through the end of our sample period.[12] Consequently, Texas presents a much different experience of the Covid-19 pandemic, relative to our three sample cities, in both the timing of its peak infection rates and in the policies enacted.

The Texas Department of State Health Services provides daily county-level counts of Covid-19 tests and positive cases. We aggregate these data by week and match them to county-level estimates of our demographic and employment data derived from the ACS. We then estimate equation (1) using

[12] In our data, statewide weekly Covid-19 infection rates in Texas started to rise the first week of June, from an average about 19 cases per 100,000 population through April and May to a high of 155 per 100,000 population for the week ending July 11. Weekly positive test rates rose from 4.7 percent for the week ending May 23 to 18.0 percent for the week ending July 11.

these data similar to before, though as with the city-level regressions, we include week rather than week × city fixed effects, and cluster standard errors by county. Also, Texas only has testing data available back to April 21, so in most cases we estimate the model from the week ending May 2 forward (i.e., the first full week we have testing data), though we have Covid-19 infection data back to March, so we estimate the unconditional relationships back to the week ending April 4. The estimates arguably present less powerful estimates of the model in equation (1) because the county-level data mask important neighborhood-level heterogeneity in employment composition, demographic composition, and infection rates. Nevertheless, the time-series behavior of the county-level estimates provide an instructive comparison to our main results.

Our estimates are shown in Figures 7 and 8. We present the results analogously to our main results in Figures 3 and 4. Our unconditional estimates show that there are essentially no differences in infection rates by type of work and no differences for counties with higher Hispanic population shares until infection rates start to rise statewide in early June. Unconditionally, counties with a higher Black population share had higher infection rates, with a 10 percentage point higher Black population share associated with about a 0.9 percentage point higher infection rate through early June. From early June forward, infection rates by type of work and by race diverged considerably. Counties with high shares of essential high-social contact workers had relatively lower infection rates in June (implying that their rates did not rise as fast as the rest of the state), though their difference dissipates by early July. Counties with higher shares of nonessential high-social contact workers saw disproportionately sharp increases in their infection rates in June and early July. Unconditionally, a 10 percentage point increase in a county's share of workers in nonessential high-social contact jobs was associated with a 0.8 percentage point higher infection rate during the week of June 6, but a 17.3 percentage point higher infection rate by the week of July 11. Over the same period, a 10 percentage point rise in a county's Black population share went from being associated with a 1.6 percentage point higher infection rate to being associated with a 2.8 percentage point higher infection rate. Similarly, a 10 percentage point rise in a county's Hispanic population share went from being associated with a 0.3 percentage point higher infection rate to being associated with a 2.8 percentage point higher infection rate. For reference, the statewide weekly infection rate rose from about 30 per 100,000 population (0.03 percent) in early June to almost 160 per 100,000 population (0.16 percent) the week ending July 11.

With some notable exceptions, adding our controls for testing and county demographic composition and jointly estimating the weekly relationships of type of work and race produce similar patterns to our main results, particularly during the periods when either citywide or statewide infection rates peak.[13] During the peak period of our main sample (April through early May), adding the controls from our full regression specification reduces the weekly coefficients on a zip code's Black population share by about 70 percent. During the peak of the Texas sample period (June through early July), adding the controls from our full regression specification reduces the weekly coefficients on a county's Black population share by just over 60 percent. Similarly, adding all controls more than accounts for the positive (unconditional) estimated weekly coefficients on a zip code's share of *nonessential* high-social contact workers during the peak of our main sample, and reduces the estimated weekly coefficients on a county's share of the same workers during the peak of the Texas sample period by about 55 percent. In both cases, much of the reduction in coefficient estimates occurs when we add the weekly interactions with local racial shares to the regression model. In contrast to our main sample, though, the coefficients on nonessential high-social contact worker shares remain large and positive throughout peak infections in Texas. Between mid-June and mid-July, a 10 percent higher share of a county's workers in nonessential high-social contact jobs is associated with a 6.1 percentage point higher

[13] As with our main estimates, in the appendix, we report the coefficient estimates for the weekly interactions with race and type of work along with their 95 percent confidence intervals. We also report the coefficient estimates for our additional controls for the fully-specified model using the Texas data.

infection rate in that county. While we cannot definitely link differences in stay-at-home policies with these differing results, it is worth noting that removal of the stay-at-home orders throughout much of Texas involved a much broader opening of nonessential businesses than in our main sample. Activities such as in-person retail shopping, indoor dining and drinking at restaurants and bars, and full patronization of personal care businesses and gyms were allowed at the start of May in Texas, but were not allowed in our sample cities before the end of June, if at all. Consequently, a broader range of nonessential workers were at work (and therefore in contact with others) in Texas during this period.

The relationship between infection rates and shares of *essential* high-social contact workers behave differently when adding controls in our main sample and in the Texas sample. Relative to the unconditional estimates, the estimates in our full specification are just over 50 percent (or 3.8 percentage points per week) lower during peak infection rates in our main sample, but 80 percent (or 1.6 percentage points per week) *higher* during peak infection rates in our Texas sample. We do not have a clear explanation for the differences between the samples, but it is worth noting that peak infection rates occurred in Texas after stay-at-home orders were lifted, while peak infection rates in our sample cities occurred concurrently with the stay-at-home orders. As a result, essential workers were not disproportionately exposed to the virus in Texas during peak infection rates, as they were in our sample cities. Interestingly, differences in testing rates during the peak periods are likely not a driver of the differential effects of our controls. Even though testing rates nationwide generally rose from March through July, the testing rates in Texas from early June through early July rise but are similar quantitatively to the testing rates in our three sample cities in April through early May. In our sample cities, weekly testing rates rose from about 100 to 400 per 100,000 population at the start of April to about 700 to 1,000 per 100,000 population at the start of May. In comparison, statewide weekly testing rates in Texas rose from 300 per 100,000 population in early June to just under 900 per 100,000 population by the week of July 11.

Finally, in our main sample, adding the controls of our full regression specification reduces the weekly coefficients on a zip code's Hispanic population share by about 55 percent during peak infection rates, but has essentially no effect on the weekly coefficients on a county's Hispanic population share during peak infection rates in Texas. We do not have a good explanation for why the controls have such a small effect on these coefficients in Texas, but we do note two caveats. First, the overall Hispanic population share in Texas is considerably higher than in our main sample (36 percent versus 27 percent). Second, given the higher population share, the fact that county-level observations mask important neighborhood-level variation in testing rates and other demographic characteristics may be more important for the Hispanic population in Texas. Once we apply all controls, however, both samples show disproportionately higher infection rates in areas with higher Hispanic population shares during the periods of peak infection rates. During the peak infection rates of our main sample, a 10 percentage point higher Hispanic population share is associated with a 1.3 percentage point higher weekly infection rate that persists for several weeks following the peak infection period. During the peak infection rates for Texas, a 10 percentage point higher Hispanic population share is associated with a 1.5 percentage point higher weekly infection rate.

Cross-Sectional Relations for Arizona and Florida Zip Codes. Finally, we compare the zip code-level relationships between infection rates, type of work, and race in the cross-section for Florida and Arizona. Like Texas, these two states had relatively low infection rates in April and most of May, and like Texas, both states had lifted their stay-at-home orders by the beginning of May to the extent that most nonessential businesses were immediately allowed to have indoor and in-person commerce. Unlike Texas, however, we do not have a time-series of Covid-19 cases for either state, nor do we have any data on testing. We only have zip code-level cumulative counts of positive Covid-19 cases for the cross-sections of zip codes from each state's public health department. Consequently, we extract these data for the July 7-8 period for each state and replicate the across-zip code correlations we reported for our main sample in Table 5. We estimate the correlations statewide and for each state's largest metropolitan area—the Phoenix Consolidated Business Statistical Area (CBSA), and the Miami-Ft. Lauderdale-Port St. Lucie Consolidated Statistical Area (CSA), respectively. For comparison, we do the same for Texas counties for the week of July 11. Our results are in Table 6.

Table 6 shows that, in general, the across-area correlations between infection rates, race, and type of work for these three states are comparable to the correlations we find for our main sample. Texas shows generally weaker correlations than our main sample, but this may be a consequence of using county-level rather than zip code-level data, since the former masks important heterogeneity across neighborhoods within counties. Where we find positive correlations, they tend to be stronger in the metropolitan areas than statewide (i.e., Phoenix shows stronger correlations than Arizona statewide and Miami-Ft. Lauderdale shows stronger correlations than Florida statewide). There are some notable exceptions for the correlations between race and employment shares. For example, the correlations between Hispanic population shares and shares of high-social contact workers are lower in Miami-Ft. Lauderdale and Florida statewide, but the correlations between Black population shares and shares of nonessential high-social contact workers are stronger in both Arizona and Florida. The correlations of type of work and race with infection rates are similar to the correlations we estimate for our main sample, especially within the Phoenix and Miami-Ft. Lauderdale metropolitan areas. Notably, the strong positive correlation between Hispanic population shares and infection rates is consistent across all areas, even across Texas counties. Thus, while we cannot identify how the time series patterns in Florida and Arizona compare to those in our main sample or in Texas, we can affirm that all of these areas show similar relationships between race, type of work, and infection rates in the cross-section. This is especially true when we focus on the major metropolitan areas in each state.

4. Conclusions

In this paper, we examine the relationship between race, type of work, and Covid-19 infection rates. Minorities have been disproportionately affected by the Covid-19 pandemic. One potential reason is that minorities tend to work in essential and high-social contact jobs, and therefore face greater exposure to the virus. We find strong heterogeneity in the residential distribution of workers in these types of jobs, with many of these workers living far from each city's central business district. We also find that these jobs are disproportionately located in neighborhoods with high minority shares and high Covid-19 infection rates. We exploit weekly variation in Covid-19 infection rates across zip codes in three U.S. cities—Chicago, New York City, and Philadelphia—to estimate their relationship to the neighborhood's employment and racial composition. We find that, unconditionally, neighborhoods with high shares of workers in high-social contact jobs, and with high shares of a Black or Hispanic population, tended to have disproportionately higher infection rates around the times when citywide infection rates peaked in our sample. Neighborhoods with high Hispanic population shares and those with high shares of residents in nonessential high-social contact jobs tended to have disproportionately high infection rates for some time after the peak. Controlling for neighborhood differences in weekly testing rates and other demographic characteristics accounts for a sizable fraction of these higher rates, but differences remain. Moreover, when we jointly estimate the weekly relationships between infection rates and type of work and infection rates and racial composition, we find that neighborhoods with high essential high-social contact worker shares or high Black population shares still exhibit higher infection rates around the time citywide infection rates peak. We also find that neighborhoods with high Hispanic population shares continue to have higher infection rates during and after the peak infection rates. In contrast, the share of nonessential high-social contact workers no longer exhibits higher infection rates during our sample period, and instead has relatively lower infection rates for much of the sample period. Thus, it is not necessarily the case that type of work accounts for the higher infection rates among Hispanics. If anything, racial composition accounts for the high infection rates we observe among some types of work. Instead, factors outside of our analysis that are specific to Hispanic neighborhoods must contribute to the disproportionately high rates observed within them. These factors may include language barriers that inhibit the transmission of vital information about the virus, local socioeconomic conditions that affect the risk of infection, or a propensity for activities that lead to higher social contact outside of work, such as greater use of public transit. To a lesser extent, factors specific to Black neighborhoods also contribute to their higher infection rates in addition to local demographic and employment differences.

We supplement our main analysis with additional evidence from Texas, Arizona, and Florida. The data are limited in the time-series and geographic variation we can exploit but provide a useful comparison to our main sample because these three states experienced a sharp rise in Covid-19 cases well after our main sample cities' rates had abated. These states also took a more aggressive approach to reopening their economies, with most of their businesses operating in at least a limited capacity by early May. Nevertheless, in the cross-section, we find similar relationships between infection rates, employment composition, and racial composition to the relationships we document using our main sample. Within Texas, where we have a time series of Covid-19 case and testing data, we find disproportionately high infection rates in areas with high minority shares or high shares of *nonessential* high-social contact workers. As with our main analysis, controlling for racial composition has a notable effect on the estimated relationships, but significantly higher infection rates remain for these groups. Notably, higher infection rates persist for nonessential high-social contact workers during the statewide peak infection rates, which coincided with the rapid reopening of these nonessential businesses. Overall, the recurring findings from our analysis are: 1) minorities are disproportionately affected by the Covid-19 pandemic, and factors like local demographic characteristics and the type of work they do only account for a portion of these effects, 2) to a lesser extent, those in high-social contact jobs are disproportionately affected by the pandemic independent of differences in race and other demographic characteristics, and 3) these effects, particularly those for type of work, are concentrated during the peak infection rates of a given location. The results suggest that periods of high infection rates amplify the risk faced by those most exposed to the virus, and that differences in demographic characteristics and type of work alone cannot explain the high infection rates Blacks and Hispanics experience during these peak periods. In ongoing work, we continue to examine the effect of other factors—such as the use of public transit, language spoken at home, and access to health insurance—and examine the relationships of race and type of work to other Covid-19 outcomes, such as the death rates and the positivity rates of tests.

References

Aaronson, Daniel, Helen Burkhardt, and R. Jason Faberman, 2020. "Potential Jobs Impacted by Covid-19," *Chicago Fed Insights Blog*, Federal Reserve Bank of Chicago, April 1.

Bartik, Alexander W., Zoe B. Cullen, Edward L. Glaeser, Michael Luca, and Christopher T. Stanton, 2020. "What Jobs are Being Done at Home During the Covid-19 Crisis? Evidence from Firm-Level Surveys," NBER Working Paper No. 27422.

Benitez, Joseph A., Charles J. Courtemanche, and Aaron Yelowitz, 2020. "Racial and Ethnic Disparities in COVID-19: Evidence from Six Large Cities," NBER Working Paper No. 27592.

Brown, Caitlin S., and Martin Ravallion, 2020. "Inequality and the Coronavirus: Socioeconomic Covariates of Behavioral Responses and Viral Outcomes Across US Counties," NBER Working Paper No. 27549.

Brynjolfsson, Erik, John J. Horton, Adam Ozimek, Daniel Rock, Garima Sharma, and Hong-Yi TuYe, 2020. "COVID-19 and Remote Work: An Early Look at US Data," NBER Working Paper No. 27344.

Chen Jarvis T., and Nancy Krieger, 2020. "Revealing the Unequal Burden of COVID-19 by Income, Race/Ethnicity, and Household Crowding: US County vs. ZIP Code Analyses," Harvard Center for Population and Development Studies Working Paper Series, 19(1)..

Dingel, Jonathan I. and Brent Neiman, 2020. "How Many Jobs Can be Done at Home?" National Bureau of Economic Research Working Paper No. 26948.

Glaeser, Edward L., Caitlin S. Gorback, and Stephen Redding, 2020. "How Much does COVID-19 Increase with Mobility? Evidence from New York and Four Other U.S. Cities." NBER Working Paper No. 27344.

Knittel, Christopher R., and Bora Ozaltun, 2020. "What Does and Does Not Correlate with COVID-19 Death Rates ," NBER Working Paper No. 27397.

Leibovici, Fernando, Ana Maria Santacreu, and Matthew Famiglietti, 2020. "Social Distancing and Contact-Intensive Occupations," *On the Economy Blog*, Federal Reserve Bank of St. Louis, March 24.

McLaren, John, 2020. "Racial Disparity in COVID-19 Deaths: Seeking Economic Roots with Census Data," NBER Working Paper No. 27407.

Mongey, Simon, Laura Pilossoph, and Alex Weinberg, 2020. "Which Workers Bear the Burden of Social Distancing Policies?" Becker-Friedman Institute Working Paper No. 2020-51.

Papageorge, Nicholas W., Matthew V. Zahn, Michèle Belot, Eline van den Broek-Altenburg, Syngjoo Choi, Julian C. Jamison, and Egon Tripodi, 2020. "Socio-Demographic Factors Associated with Self-Protecting Behavior during the Covid-19 Pandemic," NBER Working Paper No. 27378.

Table 1. Proximity Index Values and Classification by Occupation

Occupation	Proximity Index	Effective Proximity Index	Classification
Management, business, science, and arts			
Management	48.9	7.8	Low-Social Contact
Business and financial operations	49.7	10.9	
Computer and mathematical	46.1	2.3	
Architecture and engineering	50.6	25.3	
Life, physical, and social science	48.8	23.9	
Community and social service	62.1	39.1	
Legal	48.9	1.5	
Education, training, and library	59.0	10.6	
Arts, design, entertainment, sports, and media	58.7	15.8	
Healthcare practitioners and technical	84.7	80.4	
Service occupations			
Healthcare support	84.7	83.0	High-Social Contact
Protective service	70.4	66.2	
Food preparation and serving related	71.9	71.9	
Building, grounds cleaning and maintenance	53.0	53.0	
Personal care and service	77.6	63.7	
Sale and office			
Sales and related	59.1	42.6	Low-Social Contact
Office and administrative support	57.5	20.1	
Natural resources, construction, maintenance			
Farming, fishing, and forestry	44.5	44.0	High-Social Contact*
Construction and extraction	68.2	68.2	
Installation, maintenance, and repair	62.4	62.4	
Production, transportation, and material moving			
Production	56.6	56.0	High-Social Contact
Transportation and material moving	61.6	59.7	

Notes: Authors' calculations based on proximity index values from Leiboivci, Santacreu, and Famiglietti (2020) and work-from-home estimates from Dingel and Neiman (2020).

* Within mining and logging, the natural resources, construction, and maintenance occupation group is counted as low-proximity since the farming, fishing, and forestry (two-digit) occupation is the dominant occupation within this industry-occupation group pair and has a relatively low effective proximity index value.

Table 2. Essential Service Employment Shares by Industry

Industry	Essential Services Share	Effective Proximity Index	Classification
Mining and logging	0.296	47.4	Nonessential
Construction	1.000	37.5	Essential
Manufacturing	0.817	53.6	Essential
Wholesale trade	0.747	43.8	Nonessential
Retail trade	0.665	36.7	Nonessential
Transportation, warehousing, and utilities	0.994	42.1	Essential
Information	0.670	20.8	Nonessential
Finance, insurance, and real estate	0.760	23.3	Essential
Professional and business services	0.699	28.8	Nonessential
Education and Health	0.983	51.1	Essential
Leisure and hospitality	0.628	61.9	Nonessential
Other services	0.669	38.0	Nonessential
Public administration	0.980	36.2	Essential

Notes: Authors' calculations based on essential service estimates from Aaronson, Burkhardt, and Faberman (2020), proximity index values from Leiboivci, Santacreu, and Famiglietti (2020), and work-from-home estimates from Dingel and Neiman (2020). The reported effective proximity index reflects a weighted average of the proximity index values for each two-digit SOC occupation within each industry, using the 2019 industry-occupation employment shares from the Occupational Employment Statistics survey.

Table 3. Cumulative Relative Infection Rates by Race and City

	Black	Hispanic	White
Chicago	2.21	3.68	1.00
New York City	1.53	1.25	1.00
Philadelphia	2.37	1.82	1.00

Notes: Authors' calculations based on public health department data from the cities of Chicago, New York, and Philadelphia. The table reports the cumulative case rates (positive cases per 100,000 population) through July 4-7 (depending on the city) by race, normalized by the cumulative case rate for Whites.

Table 4. Racial Make-up of Employment by Type of Work

Employment Category	Percent Black	Percent Hispanic	Share of Total Employment
Essential, high-social contact jobs	16.1	24.6	21.8
Nonessential, high-social contact jobs	12.7	27.2	15.5
Essential, low-social contact jobs	11.2	11.7	35.3
Nonessential, low-social contact jobs	9.8	14.4	27.4
Total Employment	*12.2*	*17.6*	*100.0*

Notes: Authors' calculations using pooled data from the 2019 Outgoing Rotation Groups of the Current Population Survey. Sample is all employed individuals age 16 or older.

Table 5. Across-Zip Code Correlations of Employment, Race, and Infection Rates by City

I. Correlations between Racial Shares and Employment Shares				
	Chicago	**New York City**	**Philadelphia**	**Pooled Sample**
Corr(% Black, % Essential, high-contact workers)	.466	.482	.277	.454
	(.000)	(.000)	(.060)	(.000)
Corr(% Black, % Nonessential, high-contact workers)	.059	.026	.136	.056
	(.660)	(.733)	(.363)	(.349)
Corr(%Hispanic, % Essential, high-contact workers)	.419	.447	.495	.414
	(.001)	(.000)	(.000)	(.000)
Corr(%Hispanic, % Nonessential, high-contact workers)	.761	.767	.544	.719
	(.000)	(.000)	(.000)	(.000)
II. Correlations with Infection Rates (cumulative cases per 100k population through June 20)				
	Chicago	**New York City**	**Philadelphia**	**Pooled Sample**
% Essential, high-contact workers	.687	.729	.504	.630
	(.000)	(.000)	(.000)	(.000)
% Nonessential, high-contact workers	.826	.539	.168	.531
	(.000)	(.000)	(.301)	(.000)
% Black	.050	.283	.453	.126
	(.703)	(.000)	(.001)	(.034)
% Hispanic	.741	.458	-.027	.531
	(.000)	(.000)	(.858)	(.000)
N (no. of zip codes)	58	177	47	282

Notes: Authors' calculations based on data from the American Community Survey and the public health departments of the listed cities. See text for sample details. The top panel reports the across-zip code correlations between the race and employment shares listed on the left. The bottom panel reports the across-zip code correlations of each variable listed on the left with the Covid-19 infection rate, measured as the number of cumulative positive cases per 100,000 population as of July 11, 2020. All correlations are weighted by zip-code population and *p*-values are in parentheses.

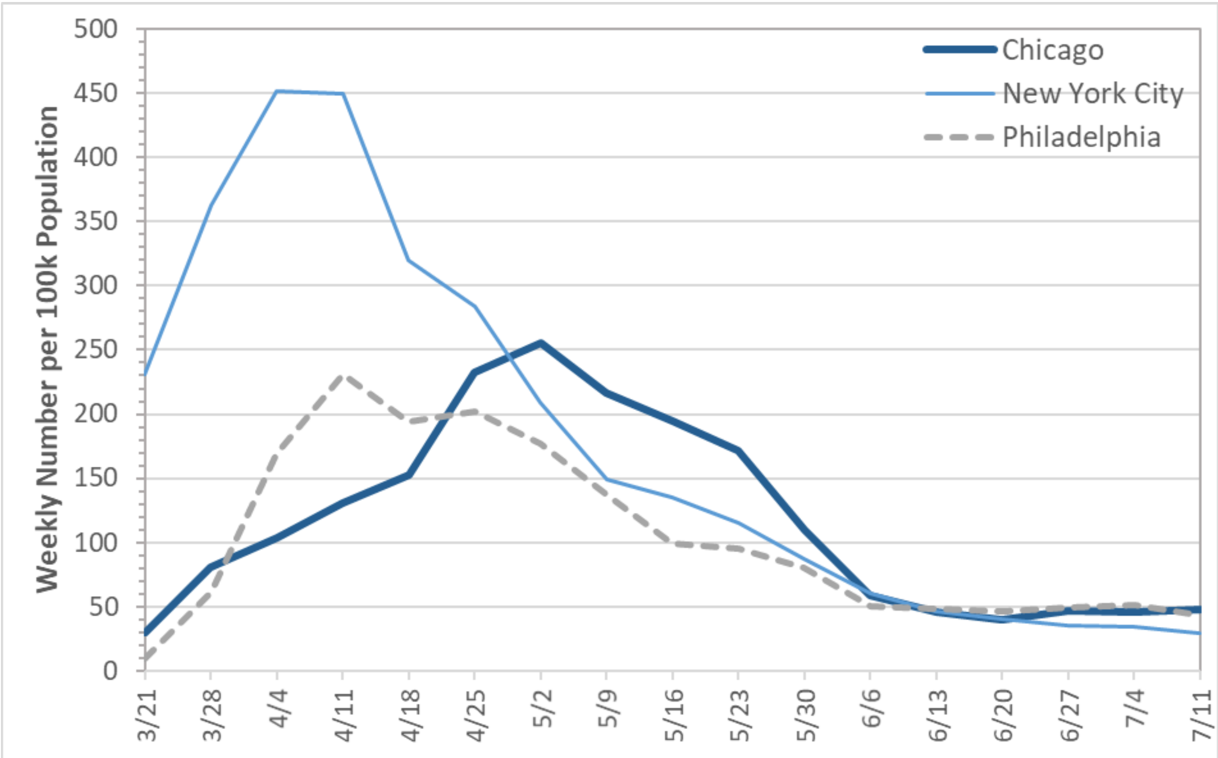
Table 6. Across-Area Correlations of Employment, Race, and Infection Rates: Selected Areas

I. Correlations between Racial Shares and Employment Shares					
	Texas, Counties	**Arizona, Zip Codes**	**Florida, Zip Codes**	**Phoenix CBSA**	**Miami-Ft. Lauderdale CSA**
Corr(% Black, % Essential,	-.044	.248	.339	.413	.440
high-contact workers)	(.481)	(.000)	(.000)	(.000)	(.000)
Corr(% Black, % Nonessential,	-.056	.219	.360	.381	.473
high- contact workers)	(.370)	(.000)	(.000)	(.000)	(.000)
Corr(%Hispanic, % Essential,	.245	.728	.158	.860	.262
high-contact workers)	(.000)	(.000)	(.000)	(.000)	(.000)
Corr(%Hispanic, % Nonessential,	.432	.601	.172	.798	-.014
high-contact workers)	(.000)	(.000)	(.000)	(.000)	(.037)
II. Correlations with Infection Rates (cumulative cases per 100k population through July 7-11)					
	Texas, Counties	**Arizona, Zip Codes**	**Florida, Zip Codes**	**Phoenix CBSA**	**Miami-Ft. Lauderdale CSA**
% Essential, high-contact	.073	.352	.237	.613	.476
workers	(.243)	(.000)	(.000)	(.000)	(.000)
% Nonessential, high-contact	.280	.319	.252	.609	.392
workers	(.000)	(.000)	(.000)	(.000)	(.000)
% Black	.212	.307	.265	.554	.143
	(.001)	(.000)	(.000)	(.000)	(.035)
% Hispanic	.402	.628	.546	.827	.441
	(.000)	(.000)	(.000)	(.000)	(.000)
N (no. of areas)	254	342	940	153	215

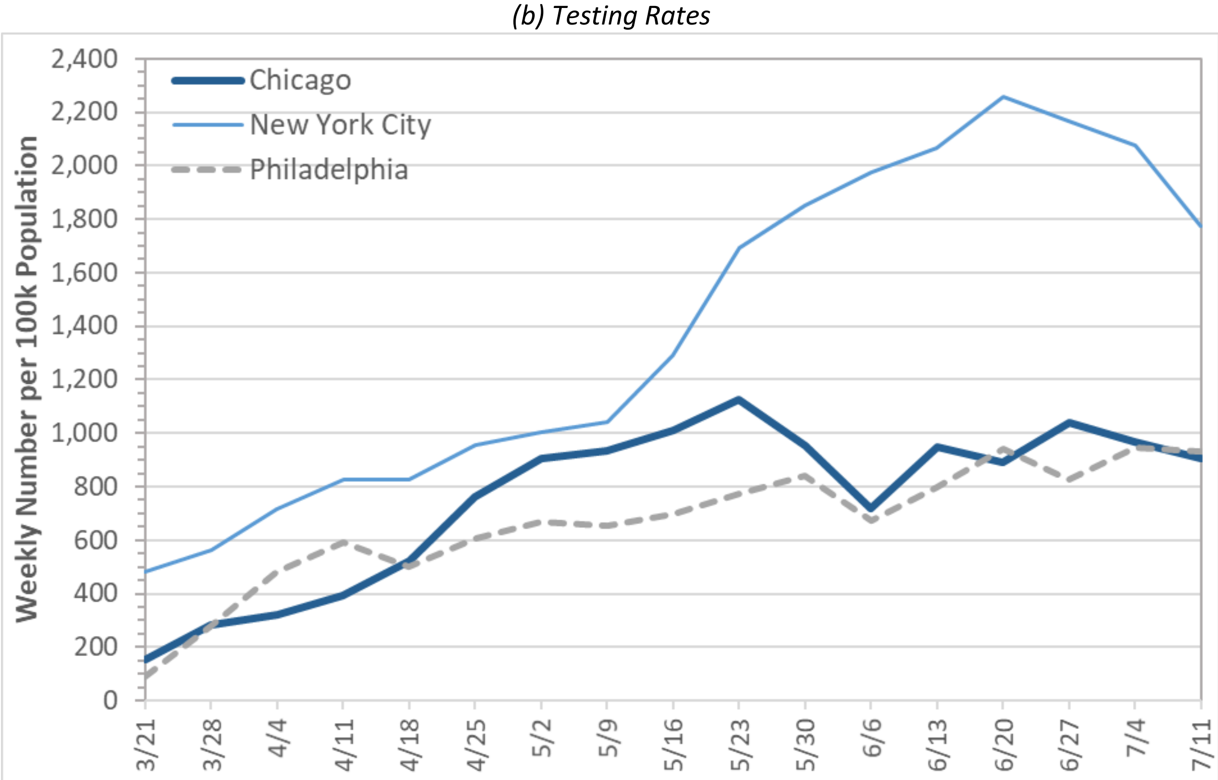
Notes: Authors' calculations based on data from the American Community Survey and the public health departments of the listed areas. See text for sample details. The top panel reports the across-county correlations (Texas) or across-zip code correlations (all others) between the race and employment shares listed on the left. The bottom panel reports the across-area correlations of each variable listed on the left with the Covid-19 infection rate, measured as the number of cumulative positive cases per 100,000 population as between July 7 and July 11, 2020. All correlations are weighted by county or zip code population and *p*-values are in parentheses.

Figure 1. Weekly Infection and Testing Rates by City

(a) Infection Rates

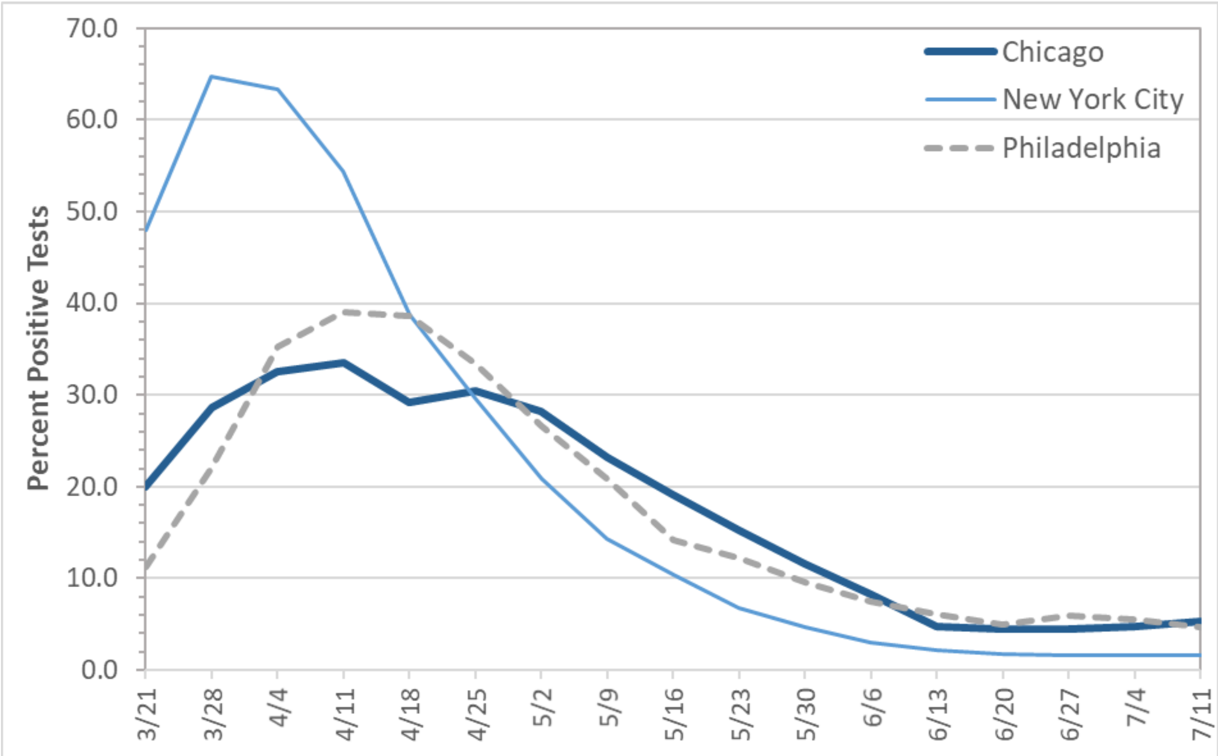


(b) Testing Rates



(continued on next page)

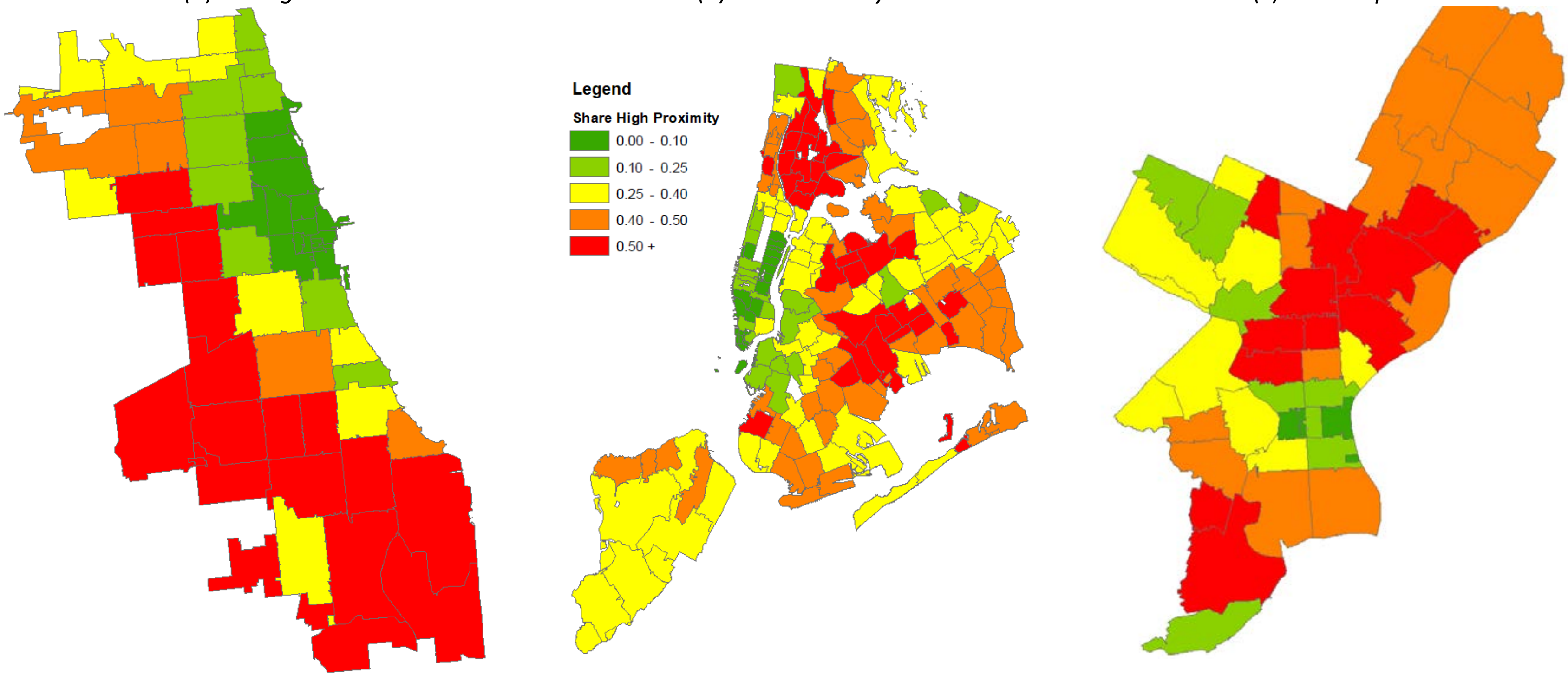
(c) Percentage of Tests that are Positive



Notes: Authors calculations based on data from the City of Chicago Department of Health and Human Services, the New York City Department of Health and Mental Hygiene, and the City of Philadelphia Department of Public Health, respectively.

Figure 2. Shares of Employment in High-Social Contact Jobs by Zip Code

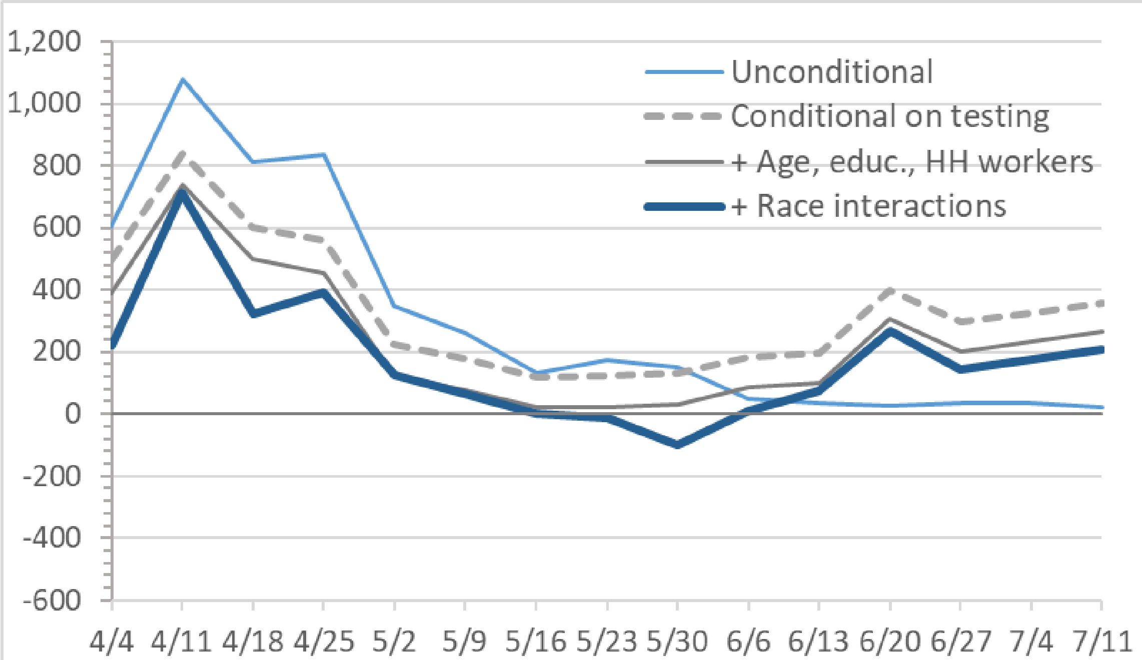
(a) Chicago *(b) New York City* *(c) Philadelphia*



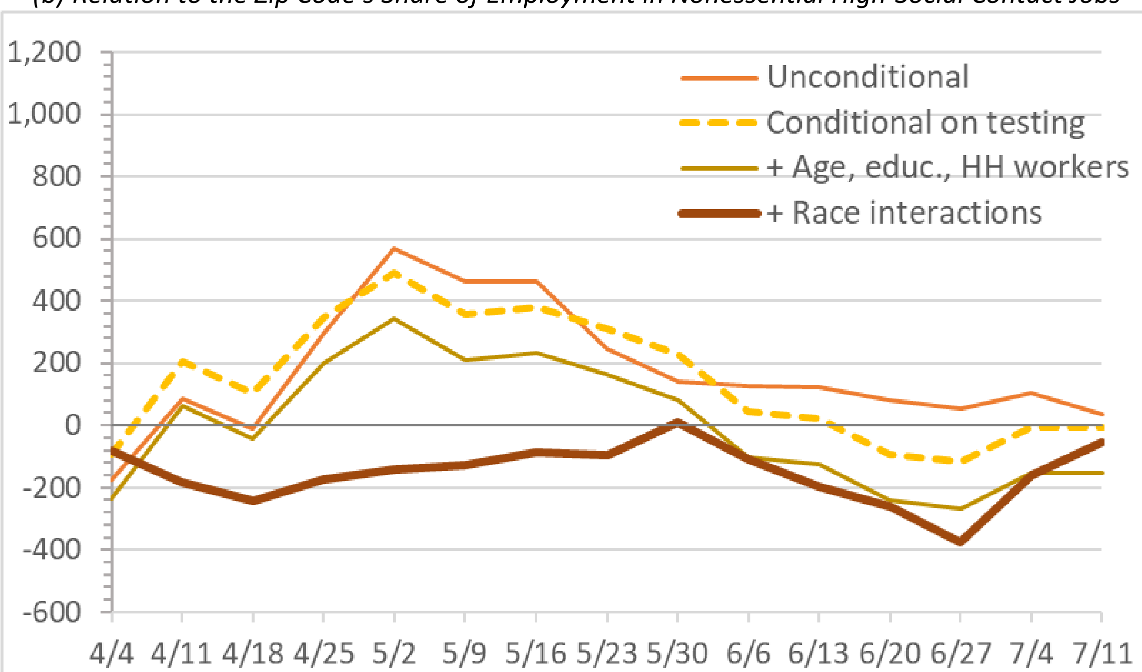
Notes: Author's calculations based on 2014-2018 industry x occupation employment data at the zip code level from the ACS. High-social contact jobs are those that require a relatively high degree personal interaction and/or have a low ability to work from home. See text and Table 1 for more details on their classification.

Figure 3. Estimates of Weekly Relations between Employment Concentration and Infection Rates

(a) Relation to the Zip Code's Share of Employment in Essential High-Social Contact Jobs



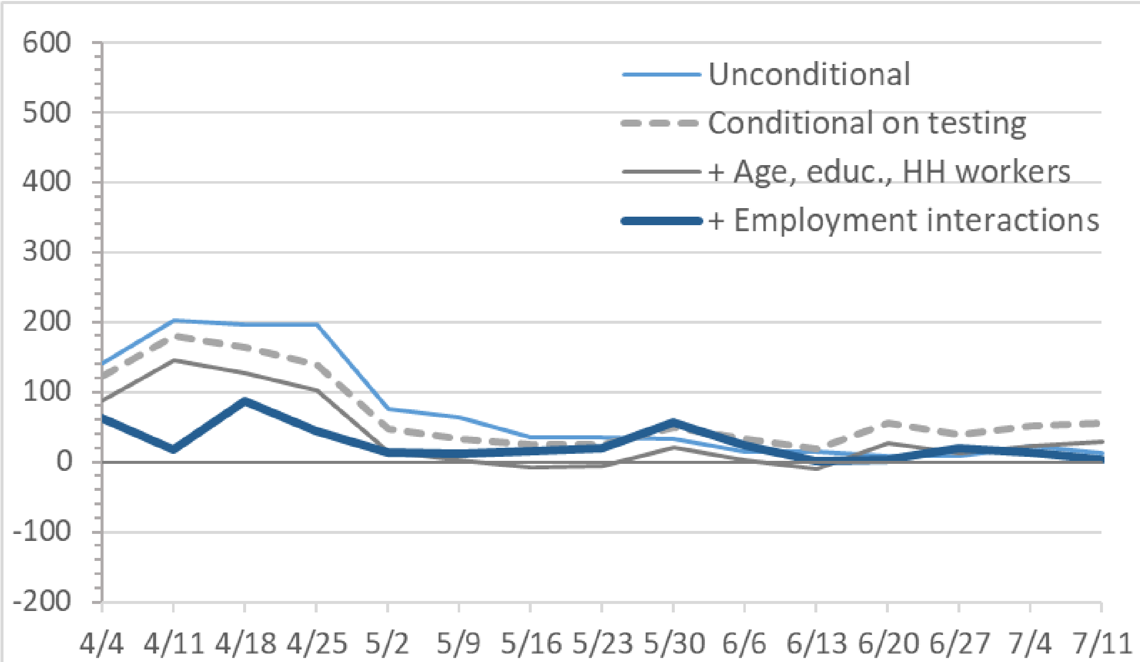
(b) Relation to the Zip Code's Share of Employment in Nonessential High-Social Contact Jobs



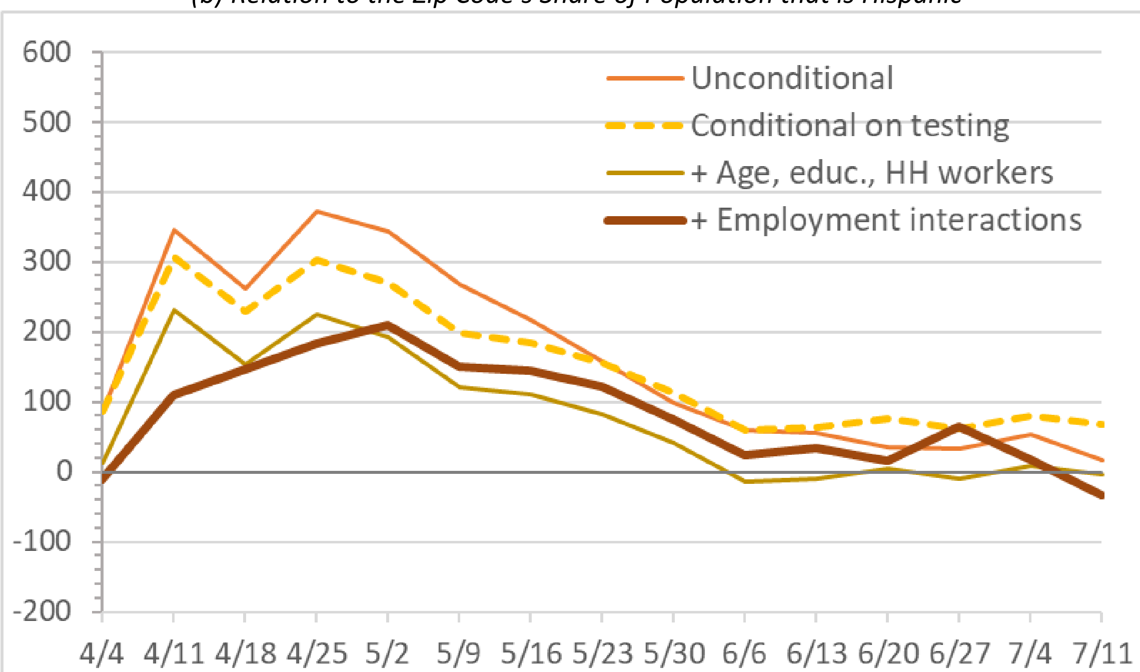
Notes: Estimates are from a panel regression of infection rates (positive cases per 100,000 population) on the week fixed effects interacted with the zip code-level employment shares reported in both panels of the figure, along with week × city interactions. The regressions are run on zip code-week observations pooled across the cities of Chicago, New York City, and Philadelphia between March 29 and July 11, 2020 (*N* = 4,030). Where listed, the regressions additionally control for weekly testing rates (tests per 100,000 population), demographic shares (population shares by age, education, and household size), and weekly interactions with racial shares at the zip code-level.

Figure 4. Estimates of Weekly Relations between Racial Composition and Infection Rates

(a) Relation to the Zip Code's Share of Population that is Black



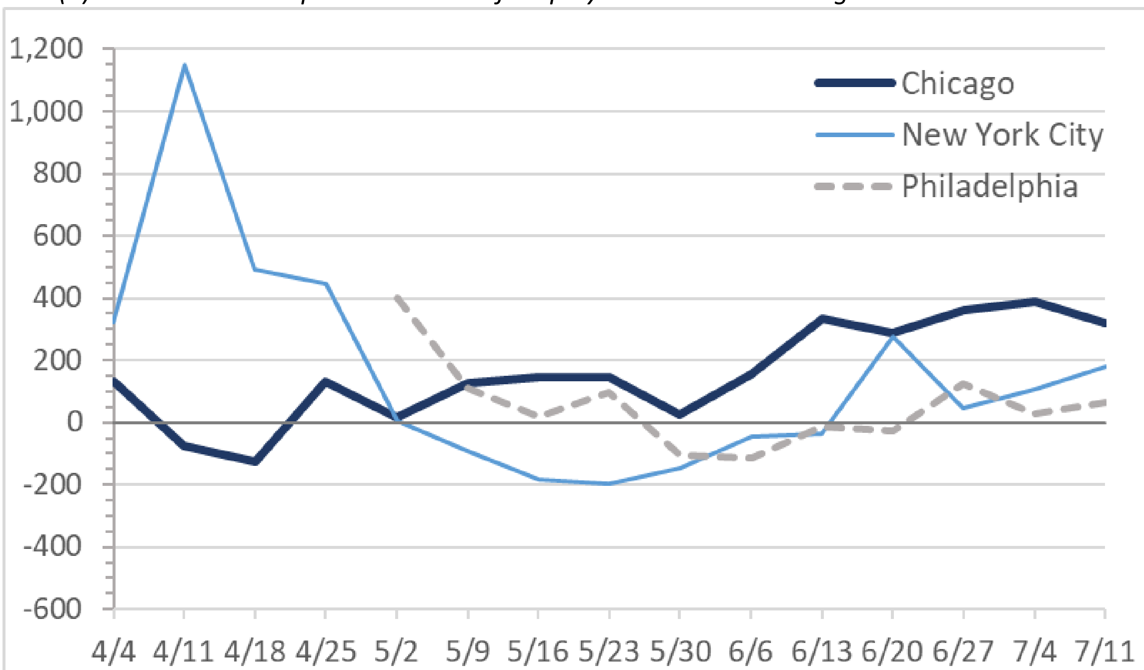
(b) Relation to the Zip Code's Share of Population that is Hispanic



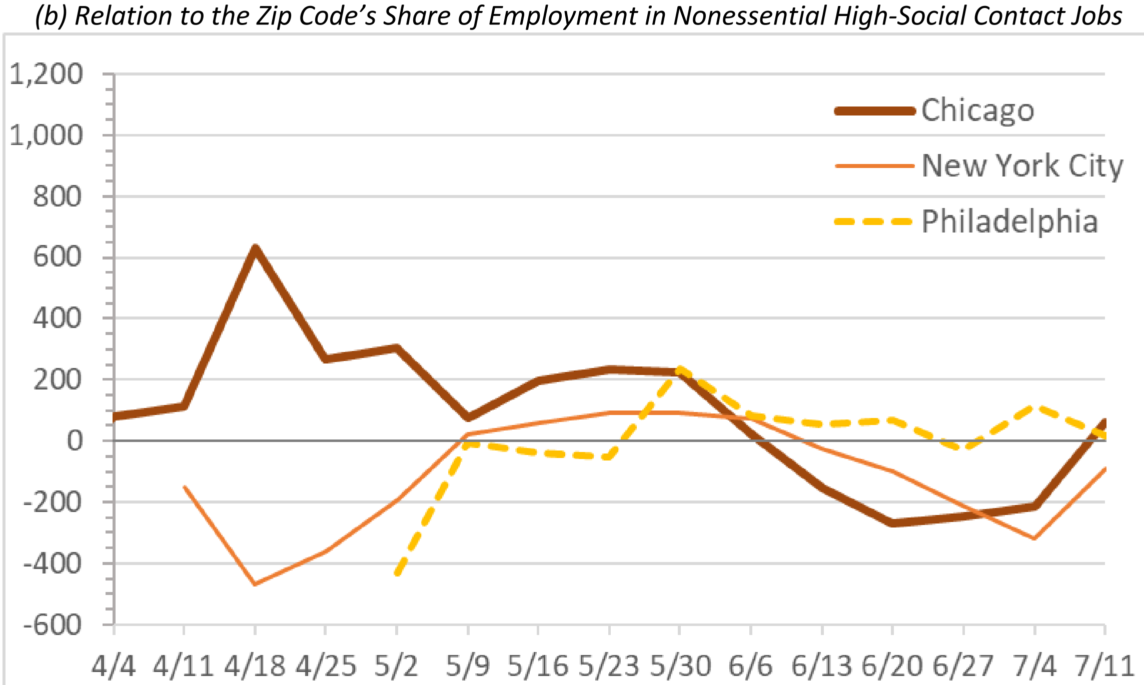
Notes: Estimates are from a panel regression of infection rates (positive cases per 100,000 population) on the week fixed effects interacted with the zip code-level race shares reported in both panels of the figure, along with week × city interactions. The regressions are run on zip code-week observations pooled across the cities of Chicago, New York City, and Philadelphia between March 29 and July 11, 2020 (*N* = 4,030). Where listed, the regressions additionally control for weekly testing rates (tests per 100,000 population), additional demographic shares (population shares by age, education, and household size), and weekly interactions with employment shares at the zip code-level.

Figure 5. Weekly Relations between Employment Concentration and Infection Rates by City

(a) Relation to the Zip Code's Share of Employment in Essential High-Social Contact Jobs



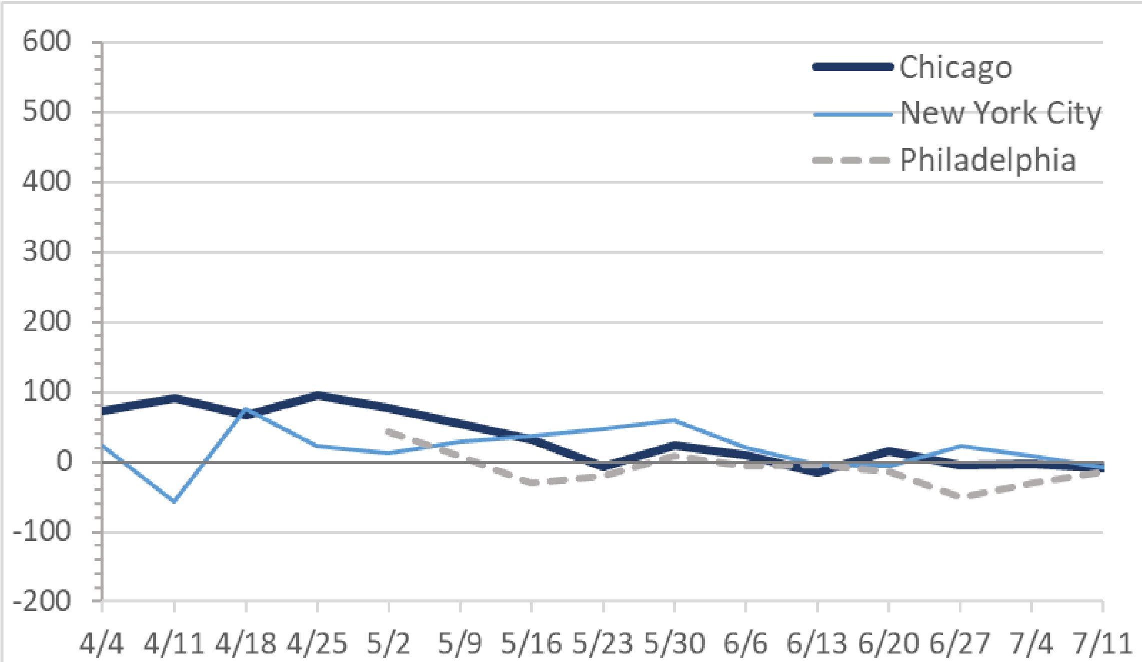
(b) Relation to the Zip Code's Share of Employment in Nonessential High-Social Contact Jobs



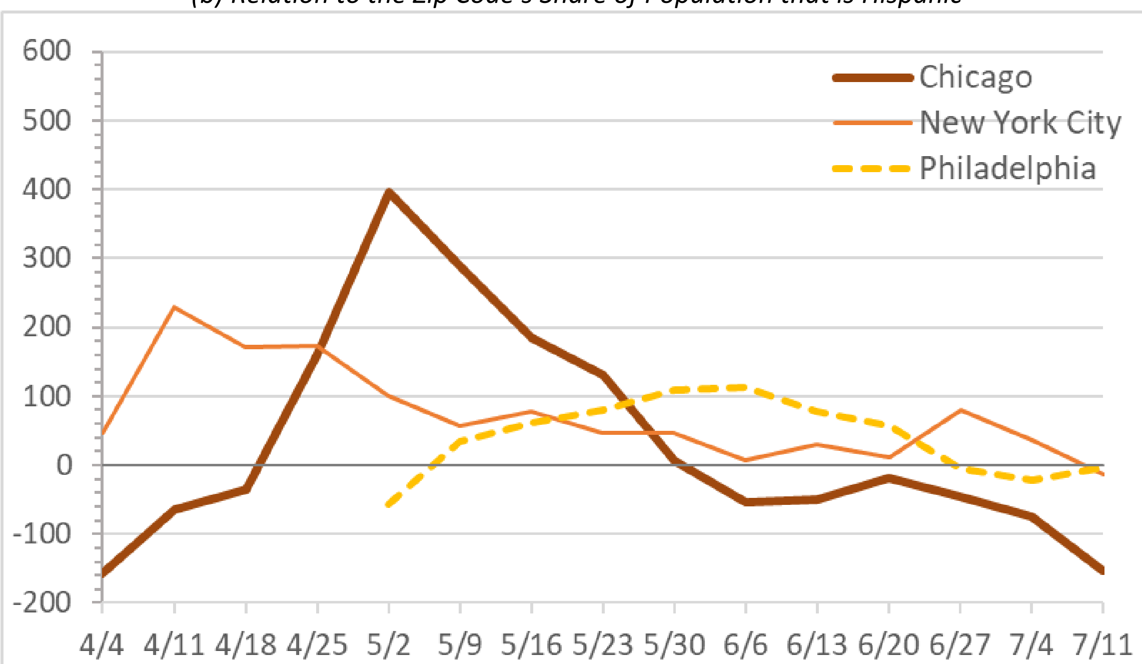
Notes: Estimates are from a panel regression of infection rates (positive cases per 100,000 population) on the week fixed effects interacted with the zip code-level employment shares (reported in both panels of the figure) and racial shares, along with week fixed effects alone, weekly testing rates (tests per 100,000 population) and additional demographic shares (population shares by age, education, and household size) at the zip code-level. The regressions are run separately on zip code-week observations for the cities of Chicago (*N* = 962), New York City (*N* = 2,655), and Philadelphia (*N* = 914) between March 29 and July 11, 2020 for weeks that each city has available data.

Figure 6. Weekly Relations between Racial Composition and Infection Rates by City

(a) Relation to the Zip Code's Share of Population that is Black



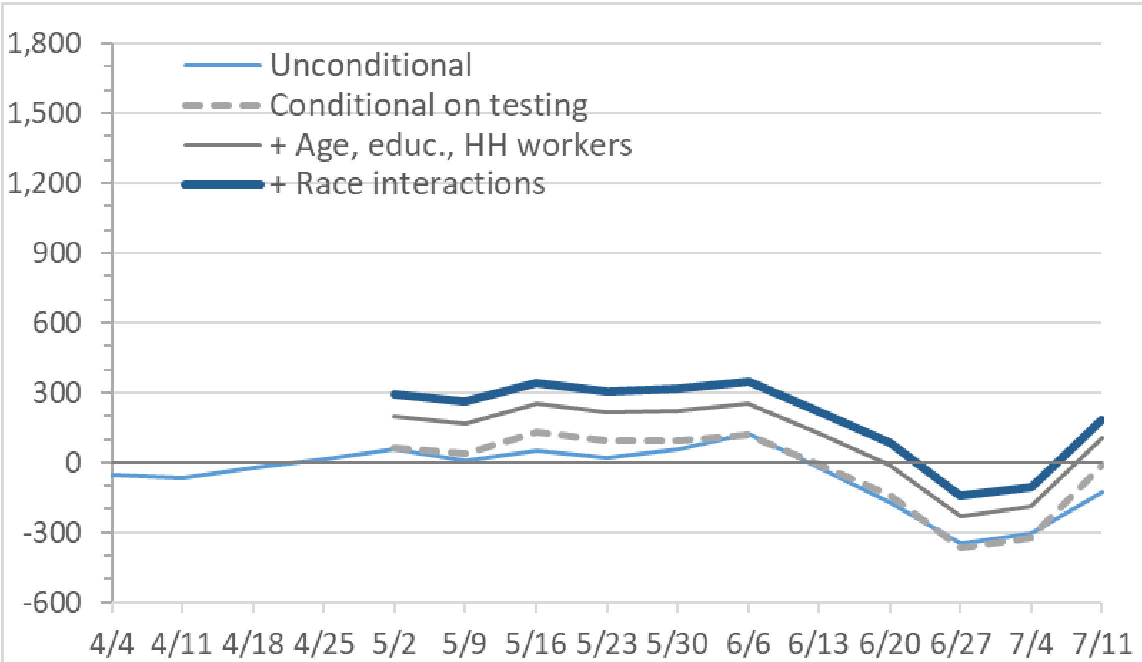
(b) Relation to the Zip Code's Share of Population that is Hispanic



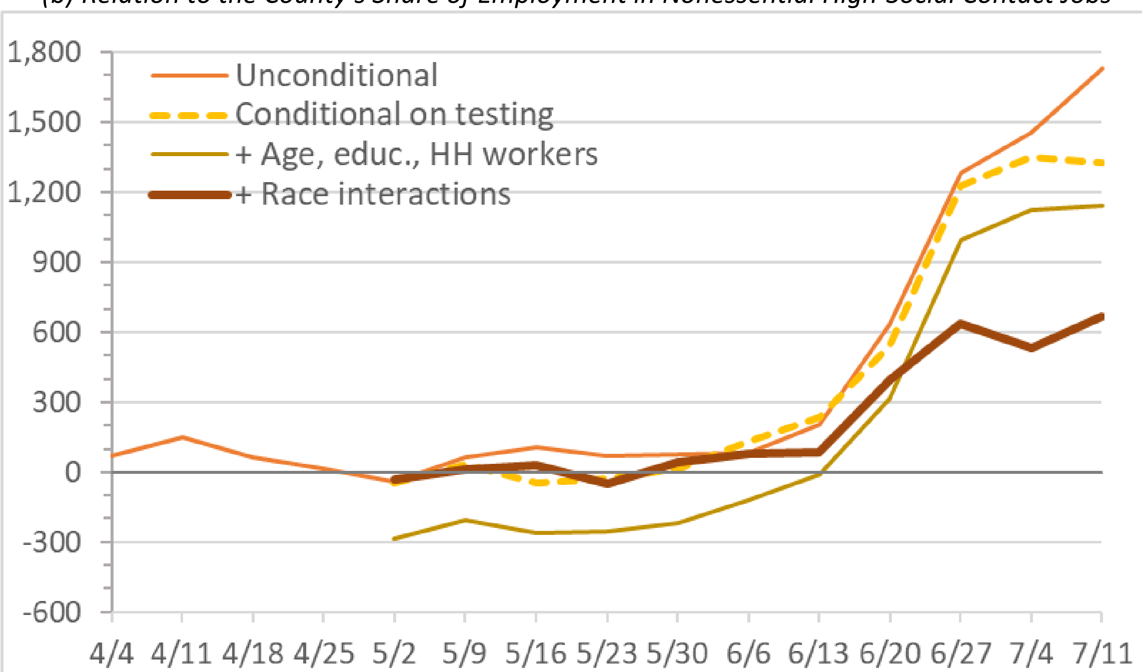
Notes: Estimates are from a panel regression of infection rates (positive cases per 100,000 population) on the week fixed effects interacted with the zip code-level race shares (reported in both panels of the figure) and employment shares, along with week fixed effects alone, weekly testing rates (tests per 100,000 population) and additional demographic shares (population shares by age, education, and household size) at the zip code-level. The regressions are run separately on zip code-week observations for the cities of Chicago (*N* = 962), New York City (*N* = 2,655), and Philadelphia (*N* = 914) between March 29 and July 11, 2020 for weeks that each city has available data.

Figure 7. Weekly Relations between Employment Concentration and Infection Rates: Texas Counties

(a) Relation to the County's Share of Employment in Essential High-Social Contact Jobs



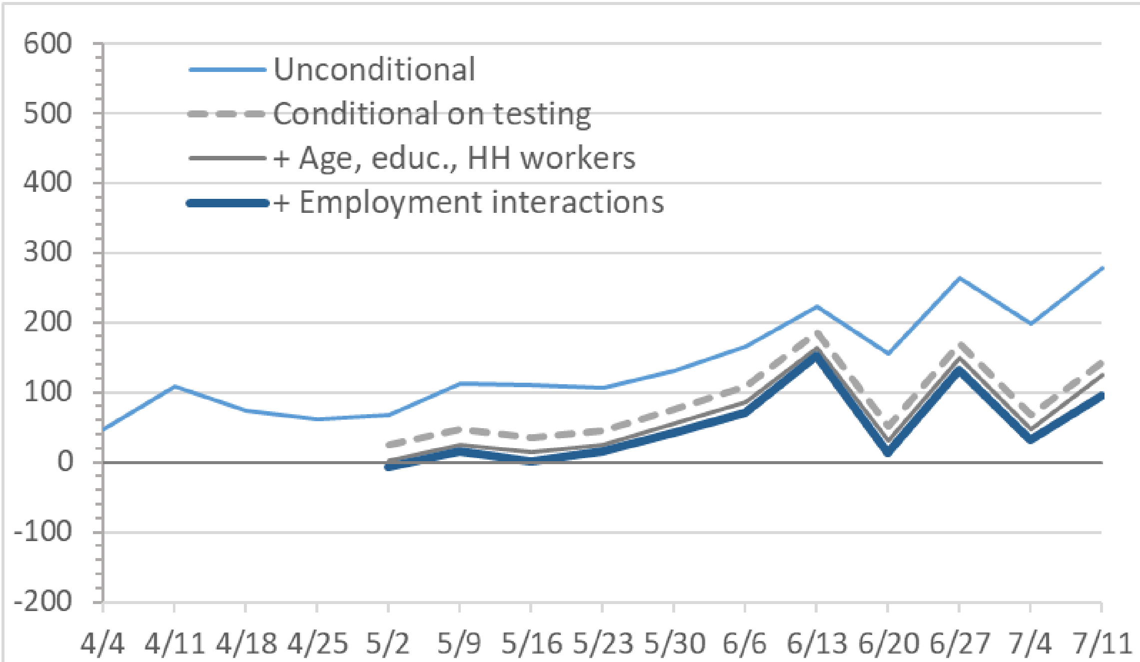
(b) Relation to the County's Share of Employment in Nonessential High-Social Contact Jobs



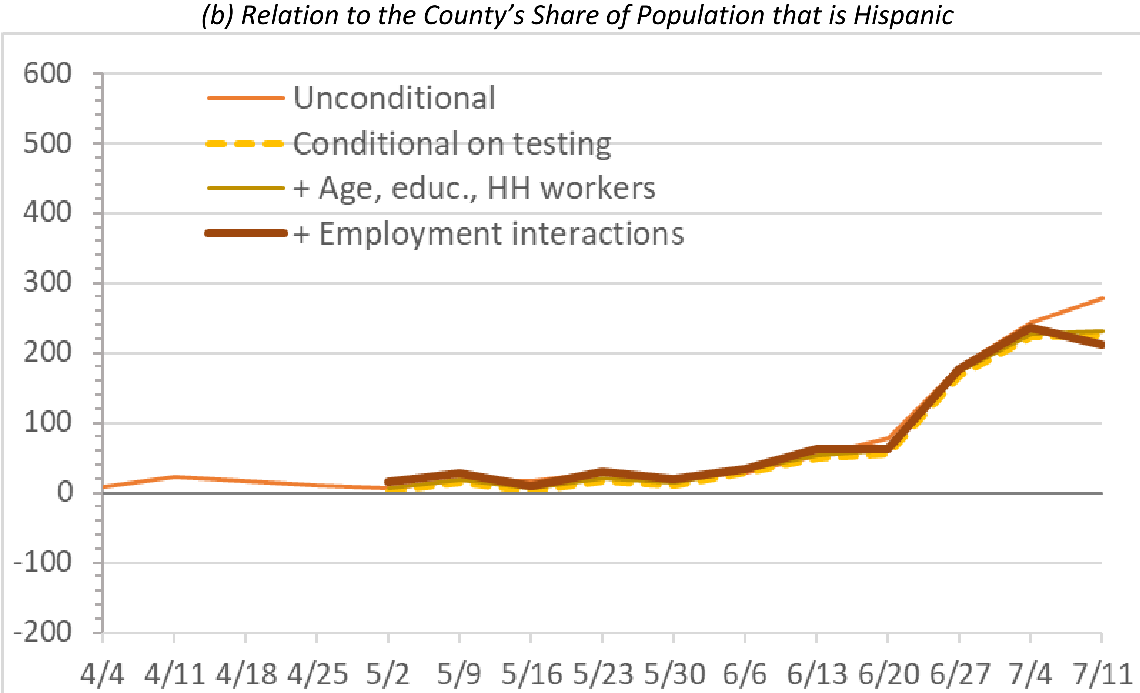
Notes: Estimates are from a panel regression of infection rates (positive cases per 100,000 population) on the week fixed effects interacted with county-level employment shares reported in both panels of the figure, along with week fixed effects. The regressions are run on county-week observations from Texas between April 25 and July 11, 2020 (N = 2,794), except for the unconditional estimates, which start on March 29 (N = 3,810). Where listed, the regressions additionally control for weekly testing rates (tests per 100,000 population), demographic shares (population shares by age, education, and household size), and weekly interactions with racial shares at the county level.

Figure 8. Weekly Relations between Racial Composition and Infection Rates: Texas Counties

(a) Relation to the County's Share of Population that is Black



(b) Relation to the County's Share of Population that is Hispanic



Notes: Estimates are from a panel regression of infection rates (positive cases per 100,000 population) on the week fixed effects interacted with the county-level race shares reported in both panels of the figure, along with week fixed effects. The regressions are run on county-week observations from Texas between April 25 and July 11, 2020 (*N* = 2,794) except for the unconditional estimates, which start on March 29 (*N* = 3,810). Where listed, the regressions additionally control for weekly testing rates (tests per 100,000 population), additional demographic shares (population shares by age, education, and household size), and weekly interactions with employment shares at the county level.

Online Appendix

This appendix reports additional results for our analysis. Figures A.1 and A.2 report the shares of each city's zip code-level shares in essential and nonessential high-social contact work, respectively. These figures separate out the shares of workers in high-social contact jobs reported in Figure 2 of the main text. The figures show that the spatial disparities in these shares persist for work in both essential and nonessential businesses, though the disparities are somewhat stronger for jobs in essential businesses.

Figure A.3 reports the estimates of the weekly relationships of type of work (top panels) and racial composition (bottom panels), when controlling for testing rates and zip code-level demographics (age, educational attainment, and household composition), and jointly estimating the weekly relationships together. The estimates are the same as those we report in Figure 3 (top panels) and Figure 4 (bottom panels) in the main text, but we report these estimates with their 95 percent confidence intervals in Figure A.3. In general, Figure A.3 shows that the relatively high Covid-19 infection rates we highlight in the text are statistically significant during their peak periods. Table A.1 reports the coefficient estimates for the testing and demographic controls included in the regressions reported in Figures 3 and 4 of the main text, along with the coefficient estimates for the full regression specification estimated using Texas counties and reported in Figures 7 and 8. The table reports the estimates along with their standard errors (clustered at the zip code or county level). Testing rates are positively related to infection rates and the estimates are highly significant, though the coefficients are notably much less than one. We find little relation of age to Covid-19 infection rates within our main sample, though counties with higher shares of prime-age individuals (age 18 to 64) have higher infection rates within Texas. We find significantly higher infection rates among areas with a higher share of residents with a high school degree or less in our main sample, but no significant differences by education in Texas.

Finally, we find that a higher share of households with at least one worker is associated with higher infection rates (regardless of the number of workers) in our main sample, but higher shares of these households have essentially no significant relation to infection rates in Texas.

Finally, Figure A.4 reports the estimates of the weekly relationships of type of work (top panels) and racial composition (bottom panels), when controlling for testing rates and zip code-level demographics (age, educational attainment, and household composition), and jointly estimating the weekly relationships together for our estimates using Texas counties. The estimates are the same as those we report in Figure 7 (top panels) and Figure 8 (bottom panels) in the main text, but we report these estimates with their 95 percent confidence intervals in Figure A.4. In general, standard error bands are wider for the Texas sample than they are for our main sample, but at their peaks, two key sets of estimates are statistically significant. These are the coefficient estimates for the county shares of nonessential high-social contact workers and Hispanic residents during statewide peak infection rates.

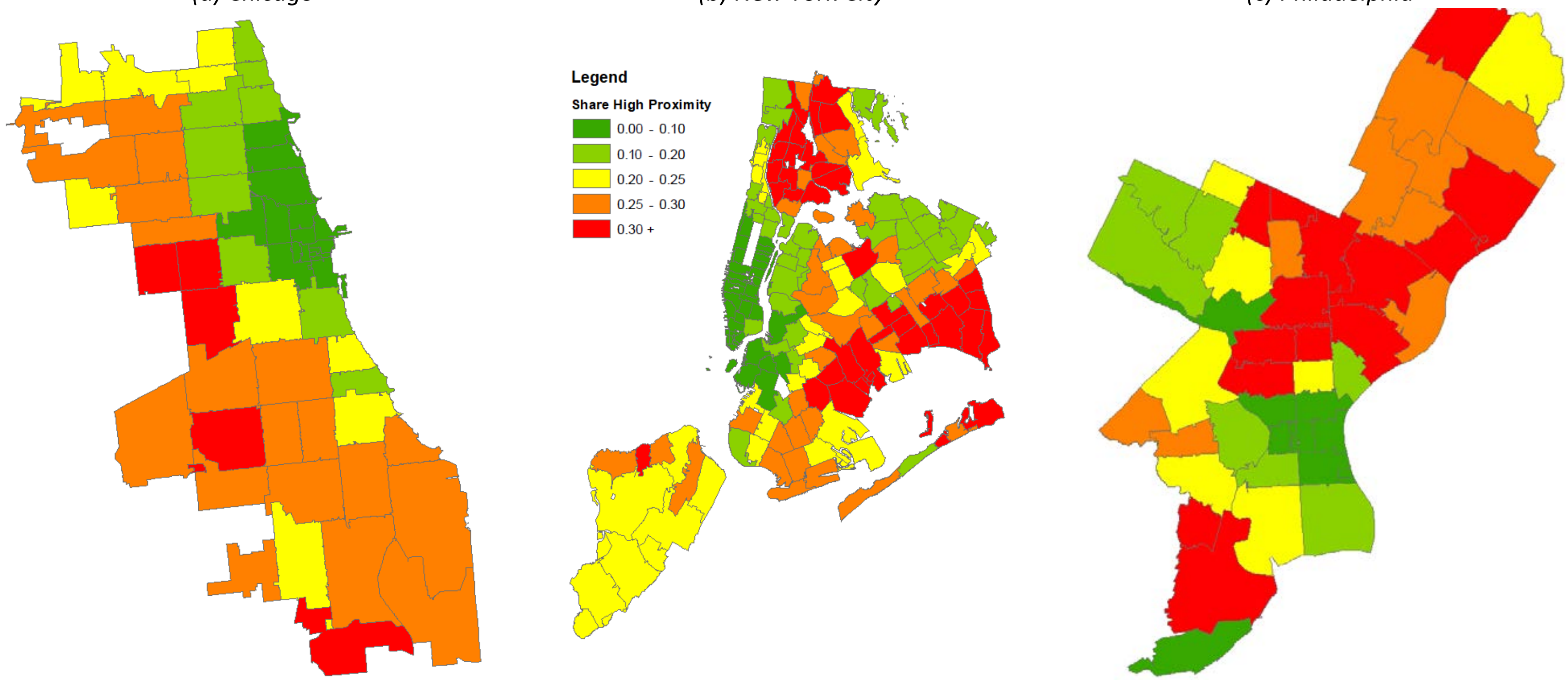
Table A.1. Coefficient Estimates for Additional Controls in Infection Rate Regressions

	Pooled City Sample			Texas
	(1)	(2)	(3)	(3)
Testing Rate (tests per 100k population)	0.131	0.130	0.127	0.038
	(.018)	(.015)	(.016)	(.008)
Share age 18 to 39	-59.6	-150.8	-64.9	404.1
	(48.5)	(42.1)	(41.8)	(127.1)
Share age 40 to 64	-97.9	-129.2	-58.5	312.2
	(76.7)	(69.3)	(60.0)	(191.5)
Share age 65 or more	70.7	103.8	154.6	249.1
	(65.7)	(67.7)	(65.3)	(179.9)
Share with HS degree or less	115.0	119.8	127.5	-75.9
	(42.8)	(24.0)	(36.2)	(56.2)
Share with some college	6.3	78.6	4.3	-74.9
	(54.9)	(52.7)	(54.4)	(73.1)
Share with one worker in household	168.8	155.4	149.0	54.0
	(50.5)	(51.6)	(45.7)	(128.5)
Share with two workers in household	91.6	115.2	116.6	121.8
	(49.5)	(49.1)	(43.4)	(130.2)
Share with three or more workers in household	244.6	325.5	277.8	-59.1
	(73.6)	(64.1)	(61.3)	(278.7)
Week x employment share interactions?	Yes	No	Yes	Yes
Week x racial composition interactions?	No	Yes	Yes	Yes
Adj. *R*-squared	0.917	0.917	0.921	0.713

Notes: Estimates are from a panel regression of infection rates (positive cases per 100,000 population) on the listed variable along with week or week × city interactions. The correspond to the weekly employment composition and racial composition interactions reported in Figures 3 and 4 (for the pooled city sample) and Figures 7 and 8 (for the Texas estimates) of the main text. The pooled city sample regressions are run on zip code-week observations pooled across the cities of Chicago, New York City, and Philadelphia between March 29 and July 11, 2020 (N = 4,030), and the Texas county sample regression is run on county-week observations between April 25 and July 11, 2020 (N = 2,794). Standard errors are reported in parentheses and are clustered at the zip code or county level.

Figure A.1. Shares of Employment in Essential High-Social Contact Jobs by Zip Code

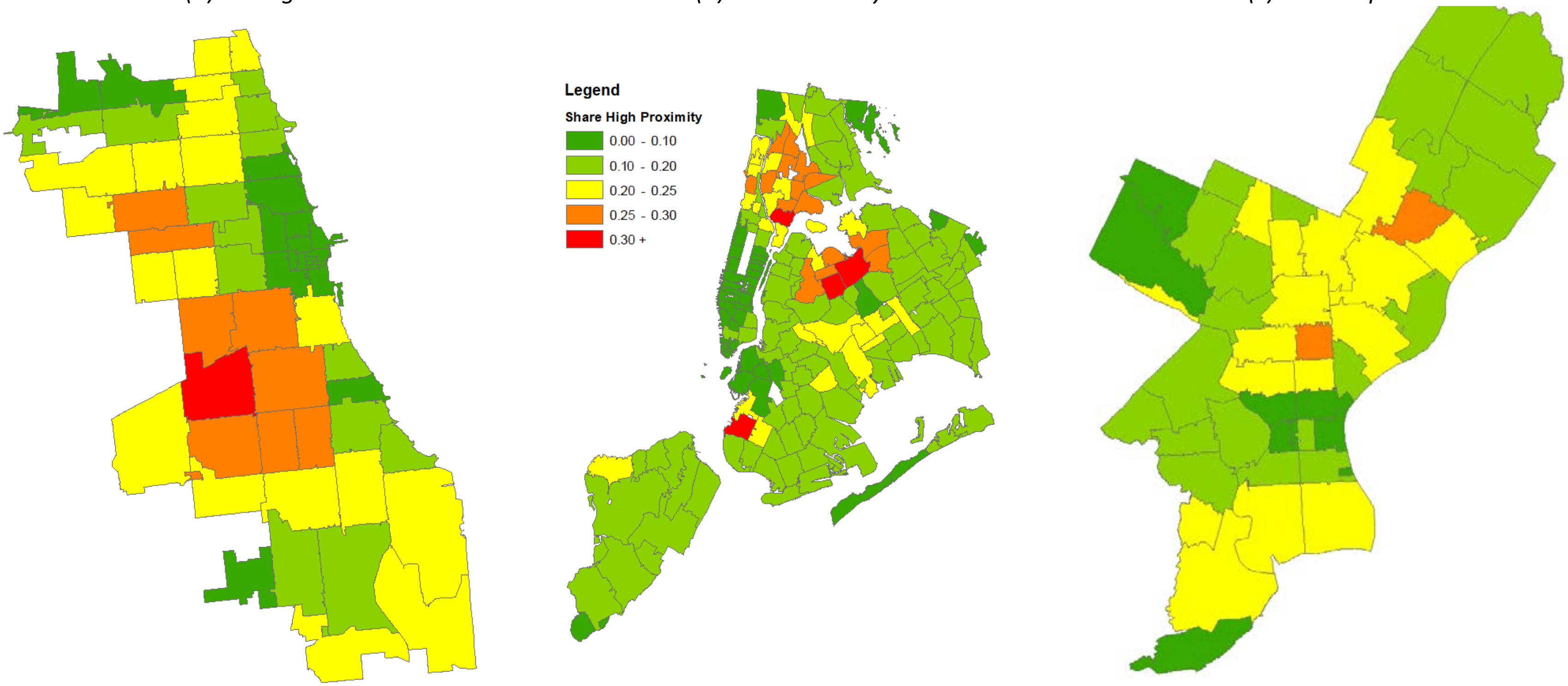
(a) Chicago *(b) New York City* *(c) Philadelphia*



Notes: Author's calculations based on 2014-2018 industry x occupation employment data at the zip code level from the ACS. Essential high-social contact jobs are those within essential businesses that require a relatively high degree personal interaction and/or have a low ability to work from home. See text and Tables 1 and 2 for more details on their classification.

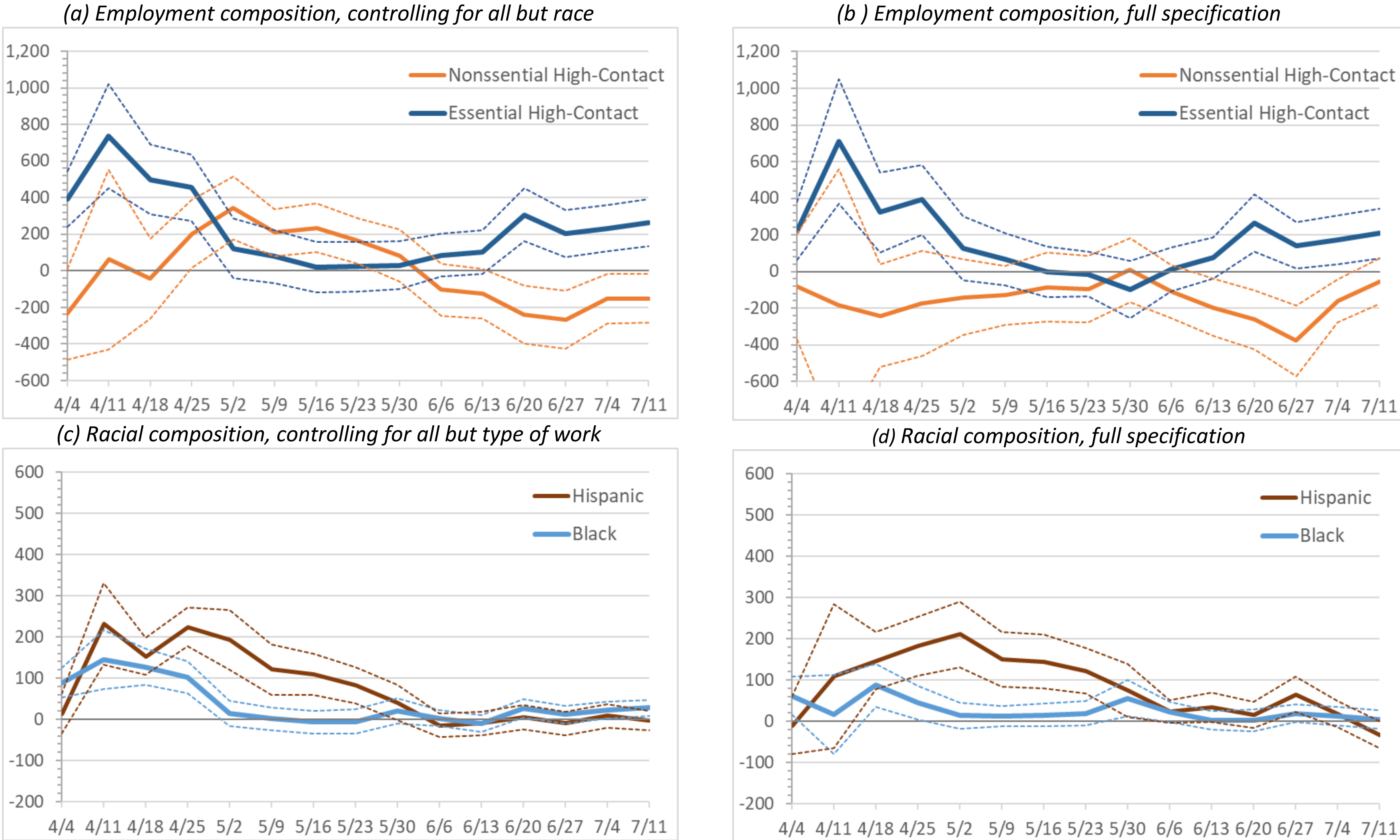
Figure A.2. Shares of Employment in Nonessential High-Social Contact Jobs by Zip Code

(a) Chicago *(b) New York City* *(c) Philadelphia*



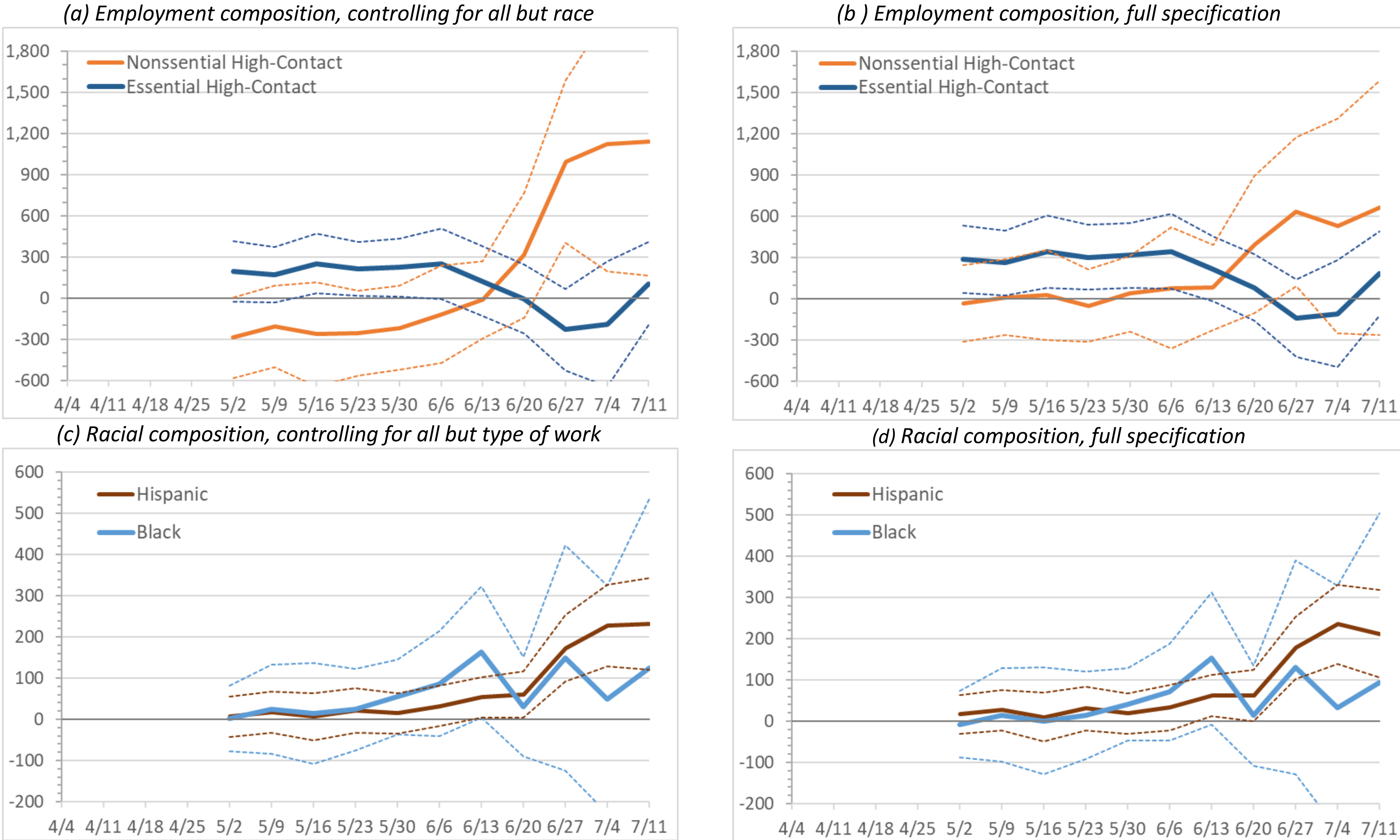
Notes: Author's calculations based on 2014-2018 industry x occupation employment data at the zip code level from the ACS. Nonessential high-social contact jobs are those within nonessential businesses that require a relatively high degree personal interaction and/or have a low ability to work from home. See text and Tables 1 and 2 for more details on their classification.

Figure A.3. Estimates of Weekly Relations between Employment Composition, Racial Composition and Infection Rates, Pooled City Sample



Notes: Estimates are from a panel regression of infection rates (positive cases per 100,000 population) on the controls listed in equation (1) and reported in Figures 3 and 4 of the main text. Solid lines replicate the estimated relationships of Figure 3 (top panels) and Figure 4 (bottom panels) of the main text. Dashed lines represent 95 percent confidence intervals. Standard errors are clustered by zip code.

Figure A.4 Estimates of Weekly Relations between Employment Composition, Racial Composition and Infection Rates in Texas



Notes: Estimates are from a panel regression of infection rates (positive cases per 100,000 population) on the controls listed in equation (1) and reported in Figures 7 and 8 of the main text. Solid lines replicate the estimated relationships of Figure 7 (top panels) and Figure 8 (bottom panels) of the main text. Dashed lines represent 95 percent confidence intervals. Standard errors are clustered by county.