Inequality and Social Distancing during the Pandemic

Caitlin Brown and Martin Ravallion[1]

Abstract: We study how pre-pandemic inequalities in America influenced social distancing over the course of the COVID-19 pandemic. Richer counties tended to see more protective mobility responses in the initial (pre-pharmaceutical) phase, but less protective responses later. Near linearity of this income effect implies that inequality between counties contributed very little to overall mobility reductions. By contrast, higher within-county inequality and/or poverty measures came with substantially larger attenuations to non-residential mobility at given average incomes. There were also significant effects of the county's racial and age composition. Standard epidemiological covariates of contact rates were also relevant, controlling for the socioeconomic factors.

JEL: I14, I15, O15

Keywords: Epidemiology; COVID-19; poverty; inequality; race

Corresponding author: Martin Ravallion: mr1185@georgetown.edu

---

[1] Brown is with the Department of Economics, Manchester University, and Ravallion is with the Department of Economics, Georgetown University, Washington DC, and the National Bureau of Economic Research. The authors thank Michael Lokshin, Mead Over, Vasia Panousi and Dominique van de Walle for helpful comments.

I. Introduction

The COVID-19 pandemic in the U.S. saw a remarkable behavioral response through physical mobility. Figure 1 shows this for the U.S. by graphing mobility responses over time by type of activity.[2] The top line shows the percentage change in the time spent at home ("residential") relative to that immediately before the pandemic, while the lower lines relate to non-residential activities. We see a sharp contraction in non-residential mobility in the first half of 2020—over 40-50% declines relative to the pre-pandemic levels—which then dissipated over time as the pre-pandemic patterns started to be restored, though with smaller contractions in each of the two winters. Residential mobility had returned to something close to pre-pandemic levels by mid-2022, although mobility for work activities remained around 20% lower than just prior to the pandemic.

Whose mobility are we talking about here? An important role has been played by individual choices related to physical mobility, especially in the early stages when pharmaceutical interventions were not yet available. The individual choices comprised both personal actions (such as choosing to shop online rather than in person, and compliance with local policies) as well as efforts to influence the actions of others (advocating local "stay-at-home" mandates, for example). Those choices were undoubtedly influenced by the socioeconomic characteristics of people and communities. We can thus ask about the socioeconomic <u>incidence</u> of the mobility response.[3] Depending on that incidence, the pandemic may come to reflect, and possibly reinforce, antecedent socioeconomic inequalities.

To study the socioeconomic incidence of the dramatic behavioral responses evident in Figure 1, we merge the Google Mobility Reports across the 3,000 U.S. counties with socioeconomic characteristics as well as (more standard) covariates suggested by the epidemiological literature.[4] We use these data to try to understand the joint epidemiological and

---

[2] This is based on the Google mobility reports, as found here. Later the paper discusses these data in more detail.
[3] We use the term "incidence" to refer to how the conditional mean of the random variable of interest—in this case, the physical mobility response M—varies with the observed indicators one is conditioning on, X, allowing for random (additive, zero mean) variation in M at a given X. This can be measured by the regression function, $E(M|X)$.
[4] The classic epidemiological models of the spread of infections have assumed homogeneous populations in a given context. For example, the word "poverty" does not appear in the classic epidemiological texts by Anderson and May (1991) and Gordis (2013), nor in the latest (6th) edition of Gordis; see Celentano and Szklo (2019). There has been greater awareness of socioeconomic factors in the spread of infectious diseases with the recent emergence of the sub-field of social epidemiology (Honjo, 2004). On the need to incorporate human behavior into epidemiological models see Ferguson (2007), Fenichel et al. (2011) and Fenichel (2013).

socioeconomic covariates of the mobility responses across counties, and how these changed over the course of the pandemic.

Once we recognize that social distancing is, in part at least, an income-constrained behavioral response, new questions emerge about how exactly incomes influence those responses—questions that this paper tries to answer for the U.S. The literature has pointed to various ways that low income can curtail the ability to avoid infection through social distancing, while wealth enhances that ability. (Later we review the literature.) For example, the type of work done by richer households may be more amenable to being relocated to their homes. The (financial and physical) assets of such households may also facilitate their protection. The welfare loss to a richer household from infection may also be lower given pre-existing inequalities in health status.

These observations suggest a positive income effect on social distancing. There may also be costs of adjusting mobility, working in the opposite direction to the positive income effect. People cannot quickly and fully adjust to a substantially lower level of socioeconomic interaction when an infection threatens. More affluent areas are likely to have more social and economic interactions in both production and consumption. If the adjustment costs are high enough then we may well see less social distancing in more affluent areas.

There are also various ways that inequality within a county might be expected to influence aggregate mobility responses. Intuitively, one might expect that more unequal counties will tend to have more people who are unable to protect themselves through social distancing. However, as we will argue, this is not guaranteed to hold. Local employment effects provide an example of how higher inequality can come with more social distancing. High inequality areas tend to have more services for high income families, with those services largely supplied by workers from low and middle-income families. If, in response to the pandemic, the relatively rich within a county no longer go to restaurants or entertainment venues, or schools close, then this essentially forces the workers to also socially distance by staying at home (employed or not). Inequality can also influence the existence and enforcement of prosocial norms and local policies relevant to social distancing. The locally rich and powerful appear often to influence policy making beyond their numbers; it is an open question as to whether this would encourage or discourage social distancing.

Separately to these effects of local inequality, the large income disparities between U.S. counties—with median household income in the richest county over five times higher than the poorest—suggest that inequality between areas might play a role in aggregate social distancing. That depends on how the marginal effects on mobility of income gains vary with average incomes across counties. If a given income gain has a larger impact on local social distancing in a poorer (richer) area then the geographic inequality will reduce (increase) overall social distancing *ceteris paribus*.

Inter-county differences in income poverty can also matter. The interpretation of the effect of this covariate depends crucially on what else one controls for. Once we control for average income, the interpretation of the regression coefficient on the poverty rate changes since it then reflects relative distribution. Maybe it is not a high incidence of poverty in a county that diminishes social distancing but rather a low average income. Controlling for inequality also affects the interpretation of the poverty effect. As we will argue, controlling for both average income and income inequality still leaves important ways in which poverty can impact social distancing, including the (ironic) fact that social exclusion among some poor families may make their mobility adjustment easier.

These effects may be expected to vary over the course of the pandemic. A key factor is going to be the availability of protective pharmaceutical options to social distancing, notably vaccines. It is thus important to distinguish the pre-pharmaceutical stage from the post-pharmaceutical stage. The socioeconomic incidence may well differ.

The following section reviews the relevant literature. Section III outlines a simple theoretical model to help motivate our empirics addressing the issues raised above. Sections IV and V describe our data and econometric methods respectively. Section VI presents our empirical results for the initial, pre-pharmaceutical, phase while Section VII looks at the re-adjustment to mobility in the more recent stages of the pandemic. Section VIII concludes.

II. Insights and puzzles from the literature

America is of special interest in this context for a number of reasons. The country's high inequality combined with the fact that the cumulative COVID-19 death rate (per capita) in the U.S. was one of the highest among rich countries points to the interest in better understanding the link between inequality and protective responses to the pandemic, including social distancing.[5] Yet from what we know, we cannot say how the income distribution in the United States impacts aggregate social distancing. Framing social distancing behavior as an income-constrained choice is especially salient in the case of the U.S., where many citizens and some political leaders resisted the stricter policies found elsewhere, such as lockdowns, and relied more on essentially voluntary compliance with at most mild directives.[6]

The COVID literature has pointed to a number of socioeconomic factors with bearing on the social distancing outcomes in the U.S. While poverty has not been a prominent causative factor in traditional epidemiological models, the literature in the social sciences has pointed to ways in which poverty might be expected to increase vulnerability to infection, which would influence social distancing, albeit in constrained and ambiguous ways.

The socioeconomic inequalities in health in America at the time the pandemic arrived are well documented, with a range of health outcomes found to be correlated with income, race and education.[7] There is also evidence that the health of poorer Americans suffered more from COVID-19.[8] There are a number of possible reasons, including the economic constraints facing poorer people in social distancing. Whether people can afford to shelter-in-place is likely to depend on their employment type, job security and savings. With little or no buffer of savings to fall back on, immobility is a costly proposition for low- and middle-income workers, especially when dependent on short-term, often casual, labor. Many low-paying jobs cannot be done from home, exposing poorer families to higher contact rates (Dingel and Neiman, 2020). Papageorge et al. (2021) find that higher income is associated with more self-protective behavioral responses, with poorer workers less able to practice social distancing, and less able to tele-work. Note that there are also socially excluded (largely non-working) poor families who are (in a sense) already socially distancing.[9] Poverty has also been found to increase the odds of having diabetes and

[5] On the COVID death rates among the Group of seven rich countries see Schellekens (2022). On income inequality and poverty measures among countries of the rich world see, for example, Morelli et al (2014).

[6] Some state governments made efforts, to varying degrees, to slow the spread of the virus, such as through "stay at home" mandates. The findings of Brzezinski et al. (2020) and Gupta et al. (2020) suggest that such efforts saw impacts.

[7] See, for example, Ettner (1996), Braveman et al. (2010) and Singh et al. (2017).

[8] Chen and Krieger (2021) estimate COVID-19 death rates (per capita) that are almost twice as high for places with poverty rates over 20% as for those under 5%. Schwandt et al. (2022) find that the pandemic's negative effect on life expectancy in California was larger for areas (as defined by census tracts) with lower median household income.

[9] Survey evidence on this point (for Canada) can be found in Stewart et al. (2009).

heart disease (Gaskin et al. 2014; O'Connor and Wellenius 2012)—risk factors that would presumably encourage social distancing among vulnerable groups. Whether poor or not, those living in generally poorer areas can be expected to face tighter fiscal and administrative constraints on community and governmental effectiveness in promoting social distancing, reflecting both local revenues and the local stocks of relevant infrastructure.[10]

These arguments can be dubbed "protection effects" since they mainly operate through the greater challenges facing poorer people in protecting themselves and their communities from infections. However, this is unlikely to be the whole story. Interpersonal interactions help maintain and expand the networks that facilitate the creation and spending of wealth.[11] An income effect in the opposite direction can arise when places with higher average incomes have a higher customary (pre-pandemic) density of personal interactions in both production and consumption, including links to external sources of infection through travel (for work, shopping and leisure). Richer people may well be better connected in both work and leisure activities. In short, the costs of rapid attenuation to social and economic interactions (including with other locations where the virus is present) could well be higher in richer places. One expects these adjustment costs to be greater for some activities than others.

In addition to the potential ambiguities in the direction of the effect of higher income on social distancing, overall income inequality may matter independently of poverty. The marginal costs and benefits of extra social distancing can vary in a systematic way with income levels, such that the distribution of income matters to the aggregate responses. Suppose that larger marginal gains in social distancing from extra income are found in poorer places where residents are more vulnerable to infection; by contrast rich places enjoy adequate protection and so see lower marginal gains. Then inequality increasing, but mean-preserving, redistributions of income from poor to rich areas will reduce aggregate social distancing.

More general claims about how higher inequality impacts social distancing are hard to defend. To help fix ideas, suppose that some minimum income (y^{min}) is needed to afford social distancing. This is not a poverty line, but the analogy can be used to draw on ideas from the

[10] Brzezinski et al. (2020) discuss local governmental and community-based responses. Chiou and Tucker (2020) argue that the diffusion of high-speed internet—favouring better-off areas—encouraged more social distancing.
[11] For example, the 2018 Consumer Expenditure Survey shows that individuals in the top income quintile spend, on average, almost five times as much as people in the bottom quintile on food outside the home, and more than three times as much on entertainment. For details, see the U.S. Bureau of Labor Statistics site.

literature on poverty and inequality measurement.[12] The population share not reaching y^{min} is $F(y^{min})$ (analogous to the poverty headcount index). The mathematical relationship between $F(y^{min})$ and the Lorenz curve is given by $L'[F(y^{min})] = y^{min}/\bar{y}$ where $L(p)$ is the Lorenz curve (the income share of the poorest percentile, p, taken to be a smoothly increasing and convex function) and $\bar{y}$ is mean income. A shift in the Lorenz curve implying a higher value of some inequality index (such as the popular Gini index) can yield either a higher or lower $F(y^{min})$, depending on how exactly the curve shifts (also allowing for intersecting Lorenz curves) and what happens to the mean.[13]

The literature on social capital has bearing on understanding the scope for social distancing and how that is influenced by inequality. The key insight is that when people trust each other within communities they are more likely to attain cooperative, mutually-beneficial, outcomes.[14] Then economic inequality is relevant, to the extent that it makes it harder to attain cooperative responses.[15] However, here too, the theoretical argument is inconclusive in this context; relatively rich households in high-inequality areas may well lobby for social distancing, given the externalities involved.[16] Politics is also in play. For the U.S., Allcott et al. (2020) find that local political affiliations—whether Democratic or Republican—have influenced local policy responses to the pandemic.

Researchers have also pointed to race as a covariate of COVID-19 incidence and impacts in the U.S.[17] The racial/ethnic composition of an area can influence mobility responses; for example, Egorov et al. (2021) find greater reductions in mobility in U.S. cities with higher ethnic fractionalization. However, poverty and race are correlated in America; the official poverty rate in 2018 was 21% for Black Americans versus 12% overall (Semega et al. 2019). An independent

---

[12] A recent overview can be found in Ravallion (2022).

[13] Empirically, it is known that changes in the usual Gini index across developing countries tend to be positively correlated with changes in the absolute poverty rate but far from perfectly so, and there are cases in which they move in opposite directions. See Ravallion (2005).

[14] A large literature shows the positive impact of social capital on a range of health and economic outcomes; see, for example, Knack and Keefer (1997), Helliwell et al. (2017), and Thornton et al. (2019). In the context of infectious disease, McNeill (1977) provides examples from history.

[15] Arguments along these lines regarding the supply of public goods can be found in Bardhan et al. (2000) and Bowles and Gintis (2002).

[16] Starting from a situation of unequal power and no enforcement of cooperative norms, enforcement by the more powerful agents can still emerge in equilibrium, as demonstrated by Anderlini et al. (2022).

[17] Brodeur et al. (2021) review the literature. For evidence on the gradient by race in COVID-19 death rates see Chen and Krieger (2020). Also, see McLaren (2021) and the commentary by Yancy (2020).

effect of the racial composition on social distancing cannot be ruled out, but it remains an open question as to whether the race effect persists when we control for poverty.

Finally, the age composition may matter to social distancing responses, while also being correlated with other key covariates. It is well documented that once infected, older people are more likely to have severe symptoms leading to hospitalization, and possibly death.[18] This may induce a stronger behavioral response for this demographic, to avoid contact with others. With higher retirement rates, the elderly (or at least the non-poor elderly) will face less economic pressure to be active outside the home. Time-use surveys for the U.S. indicate that elderly people have substantially lower contact rates in normal times (Cornwell 2010). We can think of this as a lower marginal cost of extra social distancing for the elderly.

III. An expository model of social distancing

The previous section's review of the literature points to three main ways that inequality can influence the outcomes of the pandemic, via social distancing. The first relates to how behavioral responses to the viral threat were conditioned by pre-existing dimensions of inequality, including incomes; here we ask (for example) whether poorer families were less able to protect themselves. The second relates to specific ways that inequalities <u>within</u> geographic areas influence aggregate mobility and (hence) epidemiological outcomes; here we ask whether more unequal places saw more, or less, social distancing. The third is about how antecedent inequalities <u>between</u> counties influenced the outcomes for aggregate mobility responses; here the question is whether the geographic inequalities influenced aggregate social distancing.

We can bring all three together in the form of a simple model of social distancing, viewed as an economic choice. As we will see, even in this simple model, it is not obvious how inequality through any of these three channels would affect aggregate social distancing. We will emphasize two main sources of ambiguity: how the personal welfare loss from infection varies with income and how mobility in normal (pre-pandemic) times varies with income.

We do not treat mobility as a direct source of utility but rather as a requirement for pursuing desirable activities. In an epidemic that cost includes the risk of infection. The model captures the potential trade-off between the expected loss from infection and costs of adjustment

[18] On this, see CDC COVID-19 Response Team (2020) and Ioannidis et al. (2020).

to normal activities. The expected loss is the probability of infection (p) times the loss (l) if infected. (The loss is taken to be zero for the uninfected, although this can be relaxed.) Higher mobility (m) outside the home increases the contact rate and (hence) the probability of infection, but it is assumed that it does so slowly at first, when starting from a state of little or no mobility.[19] By contrast, when we get to high levels of mobility, the high contact rate means that decrements to the chance of infection can be considerable from even a small amount of social distancing. In short, we assume that the probability of infection is a strictly increasing and convex function of the chosen level of mobility.

The extent of local inequality also matters to the infection rate, through its influence on the local public response to the pandemic. We can imagine that p depends on mean mobility in the area, which depends in turn on the local distribution of income. We represent this by simply postulating a direct effect of local inequality (G) giving $p = p(m, G)$.[20] Following the discussion in Section II, we remain open as to whether p is increasing or decreasing in G.

The loss from infection is taken to be strictly positive, and a function of own-income (y). The direction of the income effect on the loss from infection is a key factor in the implications of the model. The absolute pecuniary loss may well be greater for higher-income people. However, this is a behavioral model, so the "loss" is more properly interpreted as the welfare loss. In keeping with the idea of a protection effect, the welfare loss could be expected to be larger for poorer people, while the economically well-off are able to protect themselves. Thus, our benchmark assumption is that the loss is decreasing in income. However, we recognize that this can be questioned, so we note some implications of relaxing the assumption.

There is also an adjustment cost (c), which is taken to be a strictly increasing and convex function of the extent of adjustment in response to the threat of infection, as given by the gap between the normal (pre-pandemic) level of mobility (n) and the chosen level during the pandemic.[21] We dub that gap the <u>mobility response</u>, which will be the key object for our empirical analysis. For our benchmark case, the normal level of mobility is taken to be a strictly increasing function of income although we also note some implications of relaxing that assumption. (For one special case we also assume that n is concave in y.) In general equilibrium,

[19] We treat "mobility" as a scalar, but this can be readily relaxed.
[20] "Inequality" is represented as a scalar, but this too can be relaxed.
[21] In this respect we follow standard economic models of adjustment costs as in the theory of investment decisions; see, for example, Nickell (1979, Chapter 3).

the normal activity levels can be expected to depend on the local income distribution, so we write $n = n(y, G)$.

Combining these assumptions, we can write the mobility response as:

$$M(y, G) = m(y, G) - n(y, G) \quad (1)$$

where the chosen level of mobility during the pandemic is:

$$m(y, G) = \arg\min_{m \geq 0}[p(m, G)l(y) + c(-M(y, G))] \quad (2)$$

This equates the marginal expected loss from mobility ($p_m(m, G)l(y)$) with the marginal cost of adjustment ($c'[n(.) - m(.)]$).[22] When we refer to the "income effect on mobility" we mean $m_y(.)$ while the "income effect on the mobility response" is $M_y(.)$.

Two remarks are notable. First, adjustment costs generate a continuing dependence on pre-pandemic mobility norms; absent adjustment costs, the chosen level of mobility is zero during the pre-pharmaceutical phase, fully decoupled from pre-pandemic levels. Second, as long as the probability of infection is convex in mobility ($p_{mm}(.) >0$), any upward shift in the expected loss from infection (an increase in $p_m(.)$ or $l(y)$) will reduce the preferred level of mobility. Similarly, the introduction of an effective pharmaceutical intervention that exogenously decreases $l(y)$ will increase mobility.

The income effect on the mobility response—the first of the three roles played by income inequality in influencing aggregate social distancing—is given by:[23]

$$M_y(y, G) = \frac{-p_m(.)l'(y) - p_{mm}(.)l(y)n_y(.)}{p_{mm}(.)l(y) + c''(.)} \quad (3)$$

On noting that the denominator is positive, we see that:

$$M_y(y, G) \geq (\leq)\ 0 \text{ as } \frac{-p_m(.)}{p_{mm}(.)}\tilde{l}(y) \geq (\leq)\ n_y(.) \quad (4)$$

[22] Subscripts are used to denote partial derivatives and all functions are taken to be twice differentiable. The second-order conditions for an interior optimum hold given that $p(.)$ and $c(.)$ are both convex functions of m.
[23] Here we apply the usual implicit function theorem to the first-order condition for the problem in (2), namely $p_m(m, G)l(y) = c'(n - m)$ to derive $m_y(.)$ which we then substitute into $M_y(.) = m_y(.) - n_y(.)$.

where $\tilde{l}(y) \equiv \frac{l'(y)}{l(y)}$ is the proportionate welfare loss from infection as income rises. Consider the four combinations for the sign of $l'(y)$ and $n_y(.)$. The sign of $M_y(y, G)$ cannot be determined under our benchmark assumptions, including that the (welfare) loss from infection is higher at lower incomes ($l'(y) < 0$) and that there is a positive income effect on mobility in normal (pre-pandemic) times ($n_y(.) > 0$). If the welfare loss is decreasing in income but so too is normal mobility then the mobility response will be increasing in income ($M_y(y, G) > 0$), i.e., richer counties will see less contraction in mobility in response to the pandemic. However, if the loss from infection is increasing in income ($l'(y) > 0$) and $n_y(.) > 0$ then the mobility response will be unambiguously greater (more negative) at higher incomes ($M_y(y, G) < 0$). (The sign of $M_y(y, G)$ cannot be determined in the remaining combination, with $l'(y) > 0$ and $n_y(.) < 0$.)

Turning to the second role, Equation (1) expresses how local inequality matters directly to the mobility responses, and it is readily verified that this effect is also ambiguous in this model. However, higher inequality can alter the aggregate mobility response even if it does not alter the infection rate or normal activities. That is the third role, which depends on the <u>curvature</u> of $M(y, G)$ in y.[24] A special case offers a clearer view under the extra (simplifying) assumptions that: (i) both the loss and the adjustment cost functions are linear ($l''(y) = 0$ and $c''(y) = 0$); and (ii) the infection probability function is locally quadratic ($p_{mm}(.)$ is a constant). Then, under our benchmark assumptions, the second derivative of the mobility response w.r.t. income takes the following, much simplified, form:

$$M_{yy}(y, G) = -2m_y(.)\tilde{l}(y) - n_{yy}(.) > 0 \tag{5}$$

where $m_y(.) = \frac{-p_m(.)}{p_{mm}(.)}\tilde{l}(y) > 0$. On the RHS of (5) we see two positive effects (invoking our assumption that $n(.)$ is concave). While aggregate mobility could be either increasing or decreasing in income it will be a convex function of income ($M_{yy}(.) > 0$), implying that inequality yields a lower (less negative) aggregate mobility response.

[24] In what follows we are invoking a well-known result in mathematics called Jensen's inequality, which says that if $f(x)$ is a concave function and we compare any two values x_1 and x_2 then $f(w.x_1 + (1-w)x_2) \geq w.f(x_1) + (1-w)f(x_2)$ for any $w \in (0,1)$, with the sign reversing if $f(x)$ is convex.

Our model also reveals that this is only one possible outcome. If, instead, the welfare loss is increasing in income ($\tilde{l}(y) > 0$) then M could be either convex or concave in y (though decreasing in y, as noted), implying that higher inequality has an ambiguous effect on the mobility response. This ambiguity stems from the difference in how inequality impacts aggregate mobility during the pandemic, as distinct from normal times. If we focus solely on mobility during the pandemic, then this will be reduced by higher inequality ($m(y)$ is concave in y). However, that also holds for normal activities ($n(y)$ is also concave in y). Hence the ambiguity.

IV. Data and descriptive statistics

In measuring social distancing responses, we use Google's COVID-19 Community Mobility Reports, which track changes in visits and the time spent at various activities relative to a baseline, given by the median value for the corresponding day of the week between January 3 and February 6, 2020, prior to the onset of the epidemic in the U.S.[25] Thus, we are measuring the mobility response to the pandemic, as in Equation (1). A negative (positive) value implies that mobility saw a contraction (expansion), relative to pre-pandemic levels.

The data allow us to distinguish four groups of activities, namely residential, workplaces, recreation and shopping, and public transit. These are not, of course, independent. There are signs in Figure 1 of interlinkage in these activities; in particular, workplace mobility is in large part the mirror opposite of residential mobility, while transit and shopping/recreational co-move.

We use the latest available Google mobility data at the time of writing, namely for August 2022. However, it is important to note that this includes two distinct sub-periods of the pandemic, namely the initial sharp attenuation to non-residential mobility in the first half of 2020 and the longer subsequent period of reversal back toward normal levels, albeit with smaller seasonal contractions in non-residential mobility in the two Winters, as can be seen in Figure 1.

We initially focus on the period up to mid-2020 when social distancing was a key preventative measure but there was limited (mostly state-level) effort in the U.S. to introduce or enforce social distancing, and pharmaceutical interventions (vaccines, home testing) were not yet

---

[25] The Google data are generated using the location history of Google user accounts, and are reported for groceries and pharmacy, retail and recreation, transit stations, workplaces and residential locations. The data are expected to be representative of Google users, but naturally do not include information for those who opt out of the option of tracking their location history and/or do not have regular connectivity to the Google network. More information can be found at Google's mobility website.

available. We define the pre-pharmaceutical stage (the “initial phase”) of the epidemic as its first five months, up to mid-June 2020—essentially that first sharp dip in Figure 1.

The top panel of Table 1 provides summary statistics. In the initial phase, we see a marked contraction in mobility to workplaces, transit and (less so) retail and recreational activities; only residential activities saw positive average gains. (Table 1 also provides summary statistics on mobility responses for the longer period; Section VII will examine more closely the rebound in mobility responses in the second phase.)

We use three variables to describe the distribution of household income within each county, namely median income, the Gini index of inequality (G_{ij}), and the poverty rate based on the official national poverty line. These are published by the U.S. Census Bureau, based on their household surveys.[26] Given that incomes and their distribution during the pandemic are endogenous to mobility choices, we use their pre-pandemic values, for 2018. Descriptive statistics for these and other covariates of the mobility responses are shown in the bottom panel of Table 1. Median household incomes vary widely across counties, from $25,000 per year (Wilcox County, Alabama) to $140,000 (Loudoun County, Virginia). Gini indices also vary widely, from the lowest value of 0.25 in (aptly named?) Loving County, Texas, to a high of 0.66 in East Carroll Parish, Louisiana. The high-inequality counties tend to have lower median income (r=-0.39). Overall, the poverty rate is 15%, ranging from 2.6% (Douglas County, Colorado) to 54% (Oglala Lakota County, South Dakota). The poverty rate is highly correlated with median income (r=-0.89). Note that, having controlled for the median, the poverty rate and the Gini index can be treated as measures of different aspects of inequality. They are not the same thing of course, and there is a large unexplained variance if one regresses the Gini index on the median and the poverty rate.[27] Thus, the Gini index may still have extra explanatory power for mobility, related to changes in distribution above the poverty line. As noted in Section II, the poverty rate can influence social distancing even if neither the median income nor the Gini index changes.

Population density has long been seen as a key predictor of the contact rate for an infectious disease and (hence) the spread of infection.[28] However, we suggest that the more

[26] To enhance precision, the Census Bureau augments survey-based measures using small-area estimation methods.
[27] Regressing the log of the Gini index on the log of the median and the log of the headcount index, the R^2 is 0.31.
[28] See, for example, Anderson and May (1991) and Tarwater and Martin (2001).

relevant variable is population squared per unit area. To see why, note that in a county with population N, the potential number of contacts is $\frac{N(N-1)}{2} \cong \frac{N^2}{2}$ for large N. We call N^2 per unit land area the "potential interaction density" (PID). We also allow for differing numbers of days since the first reported COVID case in the county.

Given existing evidence on race and COVID-19, we also include the population share of Black Americans.[29] Note that this may well be proxying for other unobserved factors; for example, Black Americans are more often in jobs considered essential, with greater exposure to the virus, such as health care, food preparation, and certain services.[30] Race is also correlated with PID (r=0.41), the share of the population 65 years and older (-0.30) and poverty (0.36).

We also control for the share of total voters in the county that voted for Donald Trump in the 2016 presidential election. We interpret this as an indicator of how political affiliations influence social distancing behaviour.

Finally, as a measure of social capital, we use the Chetty et al. (2022) measure of "economic connectedness" (EC) at county level. EC is defined as the share of those deemed to have low socio-economic status who have Facebook "friends" of high socio-economic status. Our expectation is that higher EC will help promote social distancing by establishing and enforcing prosocial norms locally, such as by making local employers more willing to allow flexible working arrangements for their employees. However, the Chetty et al. estimates have the drawbacks in this context that they rely heavily on imputations and that they use Facebook data from the pandemic period, raising concerns about whether they can be treated as exogenous. We use the EC numbers for robustness checks but not in the core set of covariates.

On combining the mobility data with survey-based incomes for 2018, we obtain Figure 2, which plots the mobility responses in the initial phase against pre-pandemic median household income. We see stronger social distancing through physical mobility as income rises. These

[29] The share of Black Americans refers to the proportion of the population that identifies as Black only. Results were very similar if Hispanic was included, and the covariate was re-defined as "non-white." Two possible functional forms were considered, namely the log of the population share and the log of the fractionalization index, where the fractionalization index is $2s_1(1-s_1)$ where s_1 is the population share of group 1. In nested tests, the former dominated.

[30] Recent data from the Current Population Survey show that a majority of nursing and home health aides identify as Black/African American or Hispanic/Latino, as do a majority of those working in food preparation and most personal care and service occupations; see https://www.bls.gov/cps/cpsaat11.htm

results are strongly indicative of an income-dependent mobility response to the pandemic. We will see next if this holds up when we introduce the socioeconomic/epidemiological covariates.

V. Econometric methods

Using these data we can test whether summary statistics of household income distribution within areas help explain mobility responses. We can also see if inequality between areas matters to the observed mobility responses. For the latter purpose, we use flexible nonparametric methods to study the curvature across counties in the effect of higher average income, as well as the direction of that effect—to see whether there is a systematic pattern in how marginal impacts of higher income vary with the level of income across counties.

We postulate that the mobility response for activity k—denoted M_{ij}^k for county i in state j—is given by the following function of covariates:

$$M_{ij}^k = \alpha_j^k + f^k(\mathrm{Y}_{ij}) + \beta_1^k ln\mathrm{T}_{ij} + \boldsymbol{\beta}_2^k \mathbf{X}_{ij} + \varepsilon_{ij}^k \quad (i = 1, n;\ j = 1, N; k = 1, K\ ) \qquad (6)$$

The α_j^k's are state fixed effects, to pick up any unobserved inter-state differences in (*inter alia*) state policies in response to the pandemic, though we also provide estimates without these effects (i.e., a common intercept). Y_{ij} denotes median (pre-pandemic) household income, T_{ij} is the number of days since the first case, $\mathbf{X}_{ij}$ comprises other (epidemiological and socioeconomic) covariates of social distancing, the β's are parameters to be estimated and ε_{ij}^k is an error term with a conditional mean of zero ($E\left(\varepsilon_{ij}^k \middle| \alpha_j^k, f^k(.), ln\mathrm{T}_{ij}, \mathbf{X}_{ij}\right) = 0$). Given that the effects of the postulated covariates might change over time, as the epidemic spreads, we will also allow for interaction effects between T_{ij} and other covariates.

As noted, given that the curvature of mobility adjustment relative to average income determines the contribution of inequality between counties, we are especially concerned with assuring flexibility in representing the functional form for the income term in (6). (Otherwise, our functional form assumptions may be determining the conclusion we reach on the effect of inequality on aggregate mobility responses.) In (6), we write this as some smooth, non-parametric function, $f^k(.)$, with Y_{ij} continuous, which we estimate as a partial linear regression.[31] For this purpose, we use Stata's PLREG routine (Lokshin 2006). (For some computational purposes, we also approximate $f^k(.)$ by a cubic polynomial.)

[31] For an exposition on partial linear regression see Yatchew (1998).

VI. Results for the pre-pharmaceutical period

Table 2 provides our estimates of Equation (6) by activity for the initial period of the pandemic. Figure 3 shows the non-parametric sub-functions corresponding to Table 2, i.e., the PLREG estimates of $f^k(\mathrm{Y}_{ij})$ in (6) (with state fixed effects and setting controls at mean points). In discussing the results, we will focus more on workplace mobility, though noting general patterns and important differences with other activities.

On including the extra covariates, we still find that higher median incomes are associated with greater mobility response for the non-residential activities (Figure 3). There is only modest attenuation to the income differences in mobility responses when we include the controls. At low incomes, the mobility response was small for all activities except workplaces. *Ceteris paribus*, people in higher-income counties tended to switch more out of the external activities, to spend more time at home. As can be seen in Figure 3, the effect is sizeable; for example, on going from the poorest county to the richest increases the workplace mobility response from about a 20% reduction to about 50% (relative to the pre-pandemic baseline).

The income effects are very close to linear for all except transit (Figure 3). This suggests only modest impacts of between-county inequality. To get a sense of the magnitudes, we calculated mean-preserving counterfactual mobility responses in which Y_{ij} in (6) is replaced by its national mean. (For computational convenience, we took $f^k(\mathrm{Y}_{ij})$ to be a cubic polynomial.) On subtracting this counterfactual value from the observed values, and taking the overall mean, we then have an estimate of the contribution of the between-county component of inequality. Between-county disparities in the median implied only a 0.2-0.3 (depending on specification) percentage point reduction in workplace mobility, with a roughly similar increase in residential mobility. Inequality between counties reduced transit mobility by about one percentage point. (An Addendum provides further details.)

By contrast, higher inequality within counties led to larger (more negative) attenuations to mobility outside the home, controlling for the median income and other covariates. More unequal counties were seeing more social distancing at a given median income. Based on the specification with state fixed effects, a one standard deviation increase in the (log) Gini index reduces the workplace mobility response (making it more negative) by 0.4 percentage points. Going from the lowest (log) Gini index to the highest reduces workplace mobility by 4.6

percentage points. The reduction in transit mobility is very large; again, going from the lowest Gini index to the highest reduces transit mobility by 38 percentage points. The pattern in the regression results is consistent with a mobility response to an income gain that tends to be larger at higher incomes. (In terms of the model in Section III, $M_y(.) < 0$ and $M_{yy}(.) < 0$.)

Similarly, higher (pre-pandemic) poverty incidence is associated with larger mobility responses away from home *ceteris paribus*. Again, using the specification with state fixed effects, a one standard deviation increase in the (log) poverty rate reduces the workplace mobility response (making it more negative) by 1.4 percentage points. Going from the county with the highest poverty rate to the lowest reduces workplace mobility by a sizeable 10 percentage points.

Turning to the “non-income” covariates, higher PID is associated with greater mobility adjustments (more negative values) for activities outside the home. This holds with and without state fixed effects.

We also see a strong positive effect of a higher share of elderly people on non-residential mobility responses. In other words, the curtailment to mobility was greater for those under 65. This may be surprising, given that the health of older people was more vulnerable to COVID-19. However, prior to the epidemic, the elderly were already less mobile than others, spending more time at home. So, the results are suggesting a convergence between “young” and “old” in their mobility.

We find significant effects of a higher share of Black Americans controlling for incomes and their distribution. A one standard deviation increase in the log of the share of Black Americans attenuated the contraction in workplace mobility by 1.2 percentage points. The workplace response is consistent with the expectation that the burden of essential work fell more on Black Americans.

A higher vote share for Donald Trump at county level in 2016 is associated with significantly less effort at social distancing. A one standard deviation increase in the log of the Trump vote share reduces the workplace mobility response by 2.4 percentage points, or about one third of a standard deviation. Going from the lowest vote share for Trump to the highest attenuated the mobility response for workplaces by a massive 24 percentage points.

Our results are generally robust to including the Chetty et al. (2022) measure of economic connectedness. The main exception is that the coefficient on the Gini index for workplace mobility fell in size and is only significant at the 10% level when one includes EC as a regressor. (Full details can be found in the Addendum.) This is suggestive that county-level inequality may be picking up this aspect of social capital.

To see if the effects of these covariates change systematically as the epidemic spreads over time, we re-estimated the regressions allowing interaction effects between the covariates and the number of days since the county's first case. We could not reject the null of constant effects in most cases; the main exceptions were for transit activities, for which the interaction effects were significant for many covariates, and the share of elderly, for which greater mobility reductions were observed over time. Also, the effect of inequality on workplace mobility tended to rise as the number of days since the first case rose. The Addendum provides details.

Given that these "non-income" covariates are correlated with incomes, inferences about the incidence of social distancing that ignore incomes and their distribution can be deceptive. To show just how deceptive, Table 3 gives the restricted form of Equation (6) dropping the income variables. The results suggest large biases for most of the non-income covariates. For example, the PID coefficient for workplace adjustment with state fixed effects is more than double the value in Table 2, while the coefficients for the population share of the elderly and that of Black Americans are overestimated by about 50%.

VII. Mobility readjusts over the course of the pandemic

In terms of the behavioral model in Section III, the emergence of pharmaceutical protections (notably, but not only, vaccines) reduces case fatality rates and (hence) the potential welfare losses from infection. As noted in Section III, this would increase the preferred level of mobility and reduce the mobility response. Additionally, the accumulating adjustment costs may start to tilt the balance against the protection effects in determining the socio-economic incidence. By mid-2022, the time spent at home nationally had returned to a level not much above that of the immediate pre-pandemic period (Figure 1).[32] Non-residential mobility was still attenuated but much less so.

[32] Indeed, at mid-August 2022, the change in residential mobility relative to the pre-pandemic value was almost exactly zero. Workplace mobility was still about 10% lower than its pre-pandemic level; there was a large increase in visits to parks and other recreational facilities. (Based on Google mobility reports as accessed 16 August 2022.)

We can quantify this re-adjustment to mobility during the second phase by repeating the above analysis for the latest available Google mobility data at the time of writing (August 2022). Table 4 gives the coefficients on the linear covariates for the full period, while Figure 4 gives the corresponding estimate of the non-parametric function of median income. The patterns in these results are qualitatively similar to the initial phase (comparing Tables 2 and 4, and Figures 2 and 3), but the quantitative magnitudes are generally attenuated, reflecting the rebound in mobility.

Next we breakdown the mobility change between the two phases by median income of the county. Figure 5 gives the estimated non-parametric function for the change in mobility with respect to each activity, normalized by its baseline value and controlling for other covariates. By definition, $M_{ijt}^{k}=\left(\frac{m_{ijt}^{k}}{m_{ij0}^{k}}\right)-1$ (where M_{ijt}^{k} is the mobility response for activity k, county i, state j, date t). (Recall that this is measured relative to a pre-pandemic baseline, m_{ij0}^{k}.) Going forward to an endline $t+\tau$, the dependent variable becomes $M_{ijt+\tau}^{k}-M_{ijt}^{k}=\left(\frac{m_{ijt+\tau}^{k}-m_{ijt}^{k}}{m_{ij0}^{k}}\right)$, i.e., the change in mobility normalized by the baseline value. A positive (negative) value in Figure 5 means that more (less) time was spent in that activity when compared to the initial phase of the pandemic.

For residential, we see from Figure 5 that the change is near zero for the poorest counties, but rises (in absolute value) to around a 10 percentage point decrease for the richest. The "return to work"—the workplaces component of the mobility change—was positive but fairly flat over most of the range of median incomes. There is also a large variance in mobility changes for transit locations, for which we see markedly lower (higher) mobility responses among poor (rich) counties over the period as a whole compared to the initial phase but large mobility gains at higher incomes. Recreational and shopping mobility rebounds strongly in richer counties, but less so in poor ones. The generally positive non-residential mobility changes over the course of the pandemic came with marked reductions in time spent at home and the gap rose with income—undoing the marked increase in time spent at home in richer counties in the initial phase of the pandemic.

Corresponding to Figure 5, Table 5 gives the coefficients on the other covariates for the PLREG estimates using the data for the whole period to track the change in mobility, comparing the endline in August 2022 with the initial phase.[33] We find a strongly negative (positive) effect of within-county income inequality on the changes in workplace (residential) mobility; higher inequality counties saw substantially lower workplace mobility by the endline, when compared to the initial phase. This distributional effect is absent for the poverty rate, which had a similar effect on mobility between the two phases of the pandemic.

VIII. Conclusions

We have pointed to theoretical ambiguities in the socioeconomic incidence of the mobility responses to the threat of infection. Among the socially excluded and largely non-working poor (such as the elderly or disabled poor), social distancing may not be much of a burden. Yet the cost could be high among the working poor, since such families cannot easily maintain their consumption in isolation. The pre-pandemic levels of social and economic interaction are likely to be higher for wealthier people, and they face costs of adjusting quickly to a lower level of physical mobility. The marginal effects on social distancing of income differences may also vary with income across counties, though here too we have argued that the direction of this effect of between-county inequality on social distancing could go either way.

We have then studied the mobility responses over the course of the pandemic, using county-level data for the U.S. Counties with a higher median income tended to experience greater reductions in mobility outside the home in the initial phase, though re-bounding substantially by mid-2022. This holds when we control for the poverty rate, suggesting that the effect is coming from the attenuated mobility of the non-poor in the early phase of the pandemic, alongside enhanced mobility in the subsequent re-adjustment. While pre-existing inequalities were reflected in social distancing in the pre-pharmaceutical phase, this was partly reversed later.

Counties with a higher poverty rate and higher income inequality tended to see larger declines in non-residential mobility. (Since we are controlling for average income, these are relative distributional effects, rather than absolute.) Behavioral responses through physical mobility in the pre-pharmaceutical phase were more protective of those living in more affluent <u>and</u> unequal areas. The near-linearity of the median income effect across counties implies little or no trade-off between reducing geographic inequality and the aggregate mobility response to

[33] Note that the regression coefficients in Table 5 are not simply the difference between those in Tables 2 and 4 since the PLREG estimation allows for a nonlinear (and non-parametric) sub-function.

the pandemic. The distributional effects are mainly through within-county channels, with only a small reduction in non-residential mobility attributable to between-county inequality.

Our results are not consistent with the view that higher inequality undermines social distancing, such as by eroding local prosocial norms relevant to the chances of infection. The results are more suggestive of greater scope in high-inequality areas for enforcing such norms during the pandemic, possibly backed-up by local policies. By interpretation, the same desire among the rich (buttressed by their ability-to-pay) to protect themselves from infection by personal effort helped promote local public efforts for social distancing in high-inequality areas.

We point to two broader implications of our study. First, efforts to understand social distancing, and to respond through policy, cannot ignore the distribution of income. While voluntary social distancing can be a strongly protective response, it is a response that is firmly grounded in antecedent socio-economic inequalities. The behavioral response to the threat of infection can be highly heterogeneous across income strata and over time during the course of the pandemic. However, the specific pattern of such dependence is hard to predict on *a priori* grounds, and it is not (as we have shown) simply a situation in which inequality impedes collective social distancing—indeed, our results suggest the opposite.

Second, there may be implications for social policies. Our findings suggest that, in the absence of enforced policies to support social distancing, it will be the poorer and yet relatively equal areas that are more vulnerable to the spread of infection. Self-protection is easier for those in relatively well-off and unequal areas. Our interpretation is that poorer families are less able to afford to protect themselves, which leads them to make different social-distancing choices. This suggests that there may be a role for antipoverty policies as a complement to more direct health-policy measures in combating infectious disease, especially in the initial pre-pharmaceutical phase. Such policy implications beg for exploration in greater evaluative depth than the present paper has been able to provide.

References

Allcott, Hunt, Levi Boxell, Jacob C. Conway, Matthew Gentzkow, Michael Thaler and David Y. Yang, 2020, "Polarization and Public Health: Partisan Differences in Social Distancing During the Coronavirus Pandemic," *Journal of Public Economics* 191: 104254.

Anderlini, Luca, Leonardo Felli and Michele Piccione, 2022, "The Emergence of Enforcement," Cambridge Working Papers in Economics 2250, Cambridge University.

Anderson Roy, and Robert May, 1991, *Infectious Diseases of Humans: Dynamics and Control.* Oxford: Oxford University Press.

Bardhan, Pranab, Samuel Bowles and Herbert Gintis, 2000, "Wealth Inequality, Wealth Constraints and Economic Performance," in A.B. Atkinson and F. Bourguignon (eds) *Handbook of Income Distribution Volume 1*, Amsterdam: North-Holland.

Bowles, Samuel and Herbert Gintis, 2002, "Social Capital and Community Governance," *Economic Journal* 112(483): F419-F436.

Braveman, Paula A., Catherine Cubbin, Susan Egerter, David R. Williams, and Elsie Pamuk, 2010, "Socioeconomic Disparities in Health in the United States: What the Patterns Tell Us," *American Journal of Public Health* 100(S1): S186-S196.

Brodeur, Abel, David Gray, Anik Islam and Suraiya Bhuiyan, 2021, "A Literature Review of the Economics of COVID-19," *Journal of Economic Surveys* 35: 1007–1044.

Brzezinski, Adam, Guido Deiana, Valentin Kecht, and David Van Dijcke, 2020, "The Covid-19 Pandemic: Government vs. Community Action Across the United States". INET Working Paper No. 2020-06, Oxford Martin School, Oxford University.

CDC COVID-19 Response Team, 2020, "Severe Outcomes Among Patients with Coronavirus Disease 2019 (COVID-19)," Centers for Disease Control and Prevention, United States, February 12–March 16.

Celentano, David, and Moyses Szklo, 2019, *Gordis Epidemiology Sixth Edition*. Elsevier.

Chen, Jarvis T., and Nancy Krieger, 2021, "Revealing the Unequal Burden of COVID-19 by Income, Race/Ethnicity, and Household Crowding: US County vs. ZIP Code Analyses," *Journal of Public Health Management and Practice* 27(1): S43—S56.

Chetty, Raj, Michael Stepner, Sarah Abraham, Shelby Lin, Benjamin Scuderi, Nicholas Turner, Augustin Bergeron, and David Cutler, 2016. "The Association Between Income and Life

Expectancy in the United States, 2001-2014," *Journal of the American Medical Association* 315(16): 1750-1766.

Chetty, Raj, Matthew O. Jackson, Theresa Kuchler, Johannes Stroebel, Nathaniel Hendren, Robert B. Fluegge, Sara Gong, Federico Gonzalez, Armelle Grondin, Matthew Jacob, Drew Johnston, Martin Koenen, Eduardo Laguna-Muggenburg, Florian Mudekereza, Tom Rutter, Nicolaj Thor, Wilbur Townsend, Ruby Zhang, Mike Bailey, Pablo Barberá, Monica Bhole and Nils Wernerfelt, 2022, "Social Capital I: Measurement and Associations with Economic Mobility," *Nature.* August 1.

Chin, Taylor, Rebecca Kahn, Ruoran Li, Jarvis T. Chen, Nancy Krieger, Caroline O. Buckee, Satchit Balsari and Mathew V. Kiang, 2020, "U.S. County-Level Characteristics to Inform Equitable COVID-19 Response," medRxiv.

Chiou, Lesley and Catherine Tucker, 2020, "Social Distancing, Internet Access and Inequality" *NBER Working Paper* 26982.

Cornwell, Benjamin, 2010, "Age Trends in Daily Social Contact Patterns," *Research on Aging* 33(1): 598-631.

Dingel, Jonathan and Brent Neiman, 2020, "How Many Jobs can be Done at Home?" *Journal of Public Economics* 189: 104235

Egorov, Georgy, Ruben Enikolopov, Alexey Makarin, and Maria Petrova, 2021, "Divided We Stay at Home: Social Distancing and Ethnic Diversity," *Journal of Public Economics* 194: 104328.

Ettner, Susan L., 1996, "New Evidence on the Relationship Between Income and Health," *Journal of Health Economics* 15(1): 67-85.

Fenichel, Eli P., 2013, "Economic Considerations for Social Distancing and Behavioral Based Policies During an Epidemic," *Journal of Health Economics* 32: 440-451.

Fenichel, Eli P., Carlos Castillo-Chavez, M. G. Ceddia, Gerardo Chowell, Paula A. Gonzalez Parra, Graham Hickling, Garth Holloway, Richard Horan, Benjamin Morin, Charles Perrings, Michael Springborn, Leticia Velazquez, Cristina Villalobos, 2011, "Adaptive Human Behavior in Epidemiological Models," *Proceedings of the National Academy of Sciences* 108 (15): 6306-6311.

Ferguson, Neil, 2007, "Capturing Human Behavior," *Nature* 446: 733.

Gaskin, Darrell, Roland Thorpe Jr, Emma McGinty, Kelly Bower, Charles Rhode, J Hunter Young, Thomas LaVeist, and Lisa Dubay, 2014, “Disparities in Diabetes: The Nexus of Race, Poverty, and Place,” *American Journal of Public Health* 104(11): 2147-2155.

Gordis, Leon, 2013, *Epidemiology* 5th Edition, Elsevier.

Gupta, Sumedha, Laura Montenovo, Thuy D. Nguyen, Felipe Lozano Rojas, Ian M. Schmutte, Kosali I. Simon, Bruce A. Weinberg and Coady Wing, 2020, “Effects Of Social Distancing Policy on Labor Market Outcomes,” NBER Working Paper 27280.

Helliwell, John F., Lara B. Atkin, Hugh Shiplett, Haifang Huang, Shun Wang, 2017, “Social Capital and Prosocial Behaviour as Sources of Well-Being,” *NBER Working Paper* 23761.

Honjo Kaori, 2004, “Social Epidemiology: Definition, History, and Research Examples,” *Environmental Health and Preventative Medicine* 9(5):193-199.

Ioannidis, John P.A., Catharine Axfors, and Despina G. Contopoulos-Ioannidis, 2020, “Population-level COVID-19 Mortality Risk for Non-elderly Individuals Overall and for Non-elderly Individuals Without Underlying Diseases in Pandemic Epicenters,” *Environmental Research* 188: 109890.

Knack, Stephan and Philip Keefer, 1997, “Does Social Capital Have an Economic Payoff? A Cross-Country Investigation,” *Quarterly Journal of Economics* 112(4): 1251-1288.

Lokshin, Michael, 2006, “Difference–Based Semiparametric Estimation of Partial Linear Regression Models,” *Stata Journal* 3: 377-383.

McLaren, John, 2021, “Racial Disparity in Covid-19 Deaths: Seeking Economic Roots with Census Data,” *The BE Journal of Economic Analysis & Policy* 21(3): 897—919.

McNeill, William, 1977, *Plagues and Peoples*. New York: Doubleday.

Morelli, Salvatore, Timothy Smeeding and Jeffrey Thompson, 2014, “Post-1970 Trends in Within-Country Inequality and Poverty,” in *Handbook of Income Distribution, Volume 2,* edited by Anthony B. Atkinson and Francois Bourguignon, Amsterdam: Elsevier Science.

Nickell, S.J., 1979, *The Investment Decisions of Firms*, Cambridge University Press.

O'Connor, Alane and Wellenius, Gregory, 2012. “Rural–Urban Disparities in the Prevalence of Diabetes and Coronary Heart Disease,” *Public Health* 126(10): 813-820.

Papageorge, Nicholas, Matthew Zahn, Michèle Belot, Eline van den Broek-Altenburg, Syngjoo Choi, Julien Jamison, and Egon Tripodi, 2021, "Socio-Demographic Factors Associated with Self-Protecting Behavior during the Covid-19 Pandemic," *Journal of Population Economics* 34(2): 691-738.

Ravallion, Martin, 2005, "A Poverty-Inequality Trade-Off?" *Journal of Economic Inequality* 3(2):169-182.

_____________, 2022, "Growth Elasticities of Poverty Reduction," NBER WP 30401.

Schellekens, Philip, 2022, "Death in the Group of Seven" *Pandemic: Insight*, Posted September 17.

Schwandt H., J. Currie, T. von Wachter, J. Kowarski, D. Chapman, and S.H. Woolf, 2022, "Changes in the Relationship Between Income and Life Expectancy Before and During the COVID-19 Pandemic, California, 2015-2021." *Journal of the American Medical Association,* Published online July 07, 2022.

Semega, Jessica, Melissa Kollar, John Creamer and Abinash Mohanty, 2019, "Income and Poverty in the United States: 2018," US Census Bureau.

Singh, Gopal K., Gem P. Daus, Michelle Allender, ChristineT. Ramey, Elijah K. Martin, Chrisp Perry, Andrew A. De Los Reyes and Ivy P. Vedamuthu, 2017, "Social Determinants of Health in the United States: Addressing Major Health Inequality Trends for the Nation, 1935-2016," *MPH1 International Journal of MCH and AIDS* 6(2): 139-164.

Stewart, Miriam J., Edward Makwarimba and Linda I. Reutter, 2009, "Poverty, Sense of Belonging and Experiences of Social Isolation," *Journal of Poverty* 13:173–195.

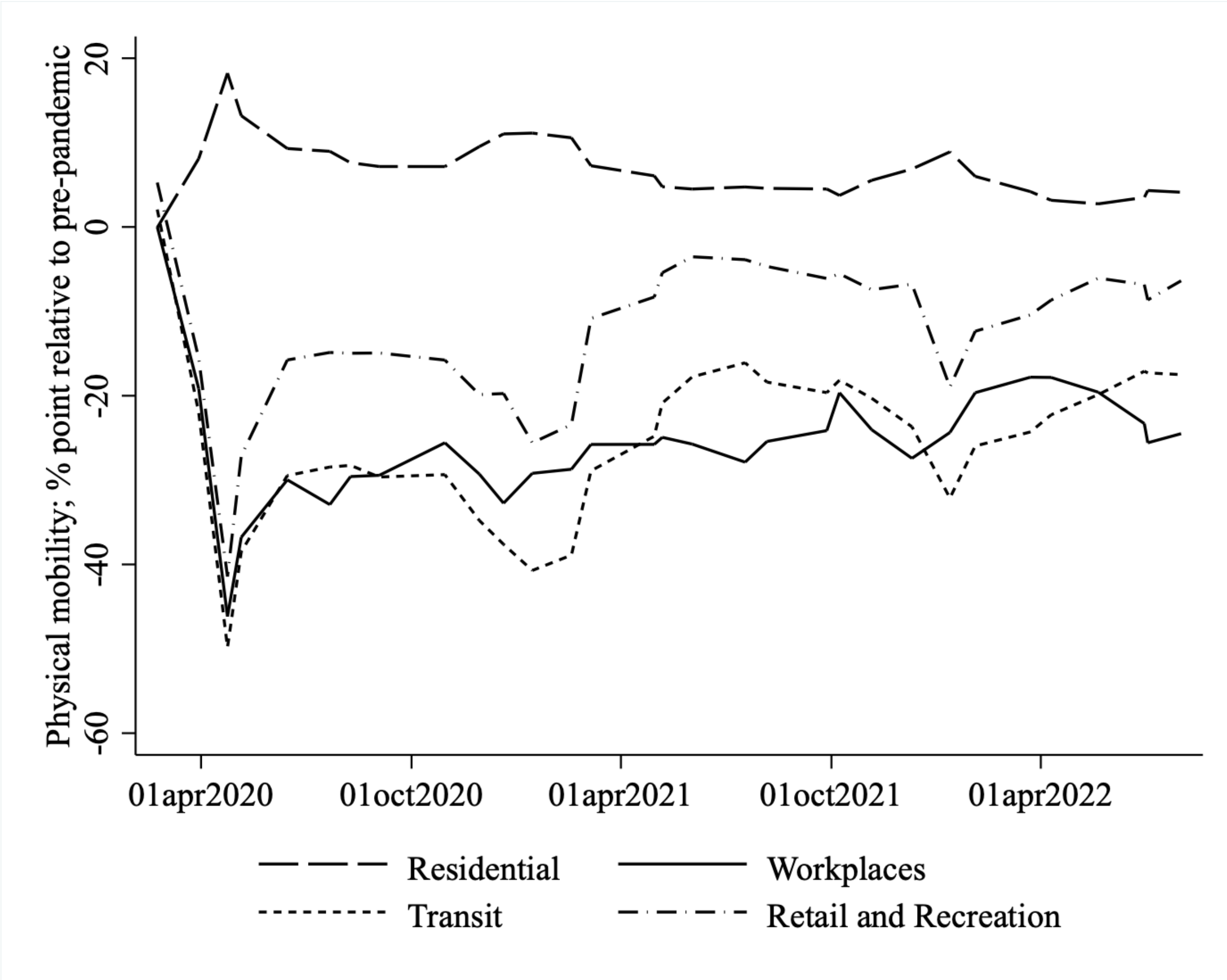
Tarwater, Patrick M., and Clyde F. Martin, 2001, "Effects of Population Density on the Spread of Disease," *Complexity* 6(6): 29-36.

Thornton, Emily M., Lara B. Aknin, Nyla R. Branscombe, John F. Helliwell, 2019, "Prosocial Perceptions of Taxation Predict Support for Taxes," *PLoS ONE* 14(11): e0225730.

Yancy, Clyde W., 2020, "COVID-19 and African Americans," *Journal of the American Medical Association* Viewpoint, April 15

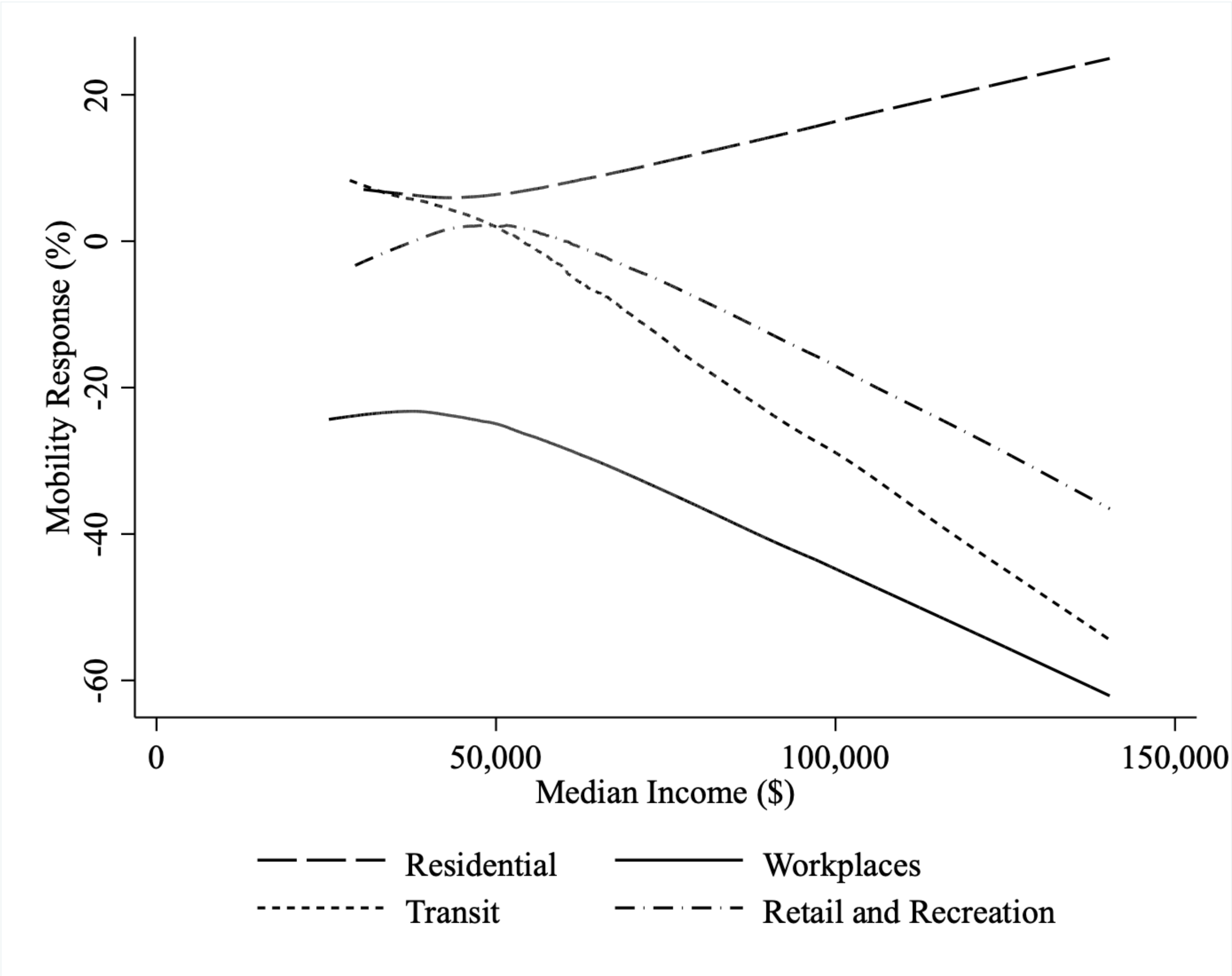
Yatchew, A., 1998, "Nonparametric Regression Techniques in Economics," *Journal of Economic Literature* 36: 669–721.

Figure 1: Changes in physical mobility in the U.S. over the course of the Covid-19 pandemic



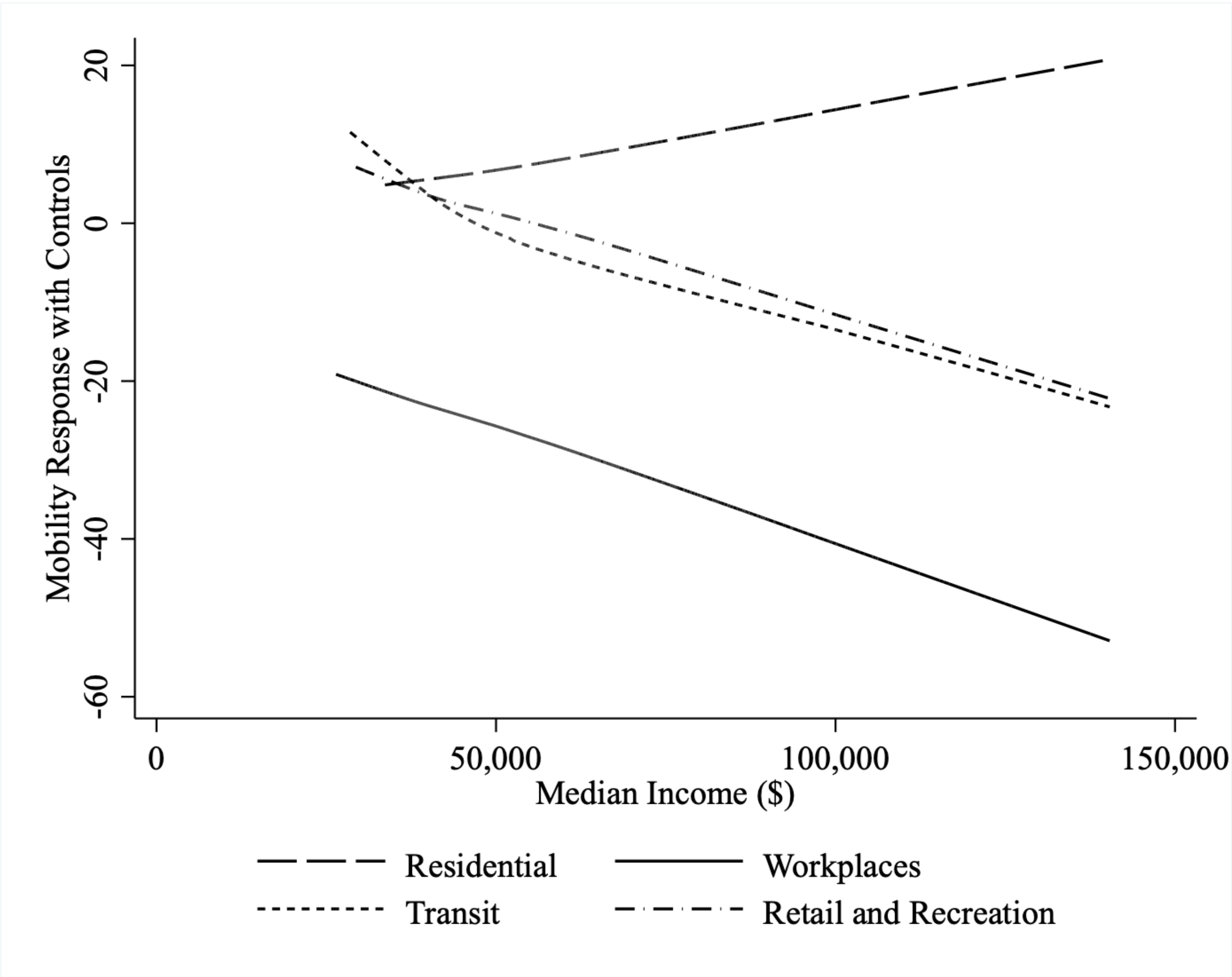
Source: Google's COVID-19 Mobility Reports up to August 8, 2022 (accessed August 16).

Figure 2: Mobility responses in the initial phase of the pandemic plotted against pre-pandemic median household income



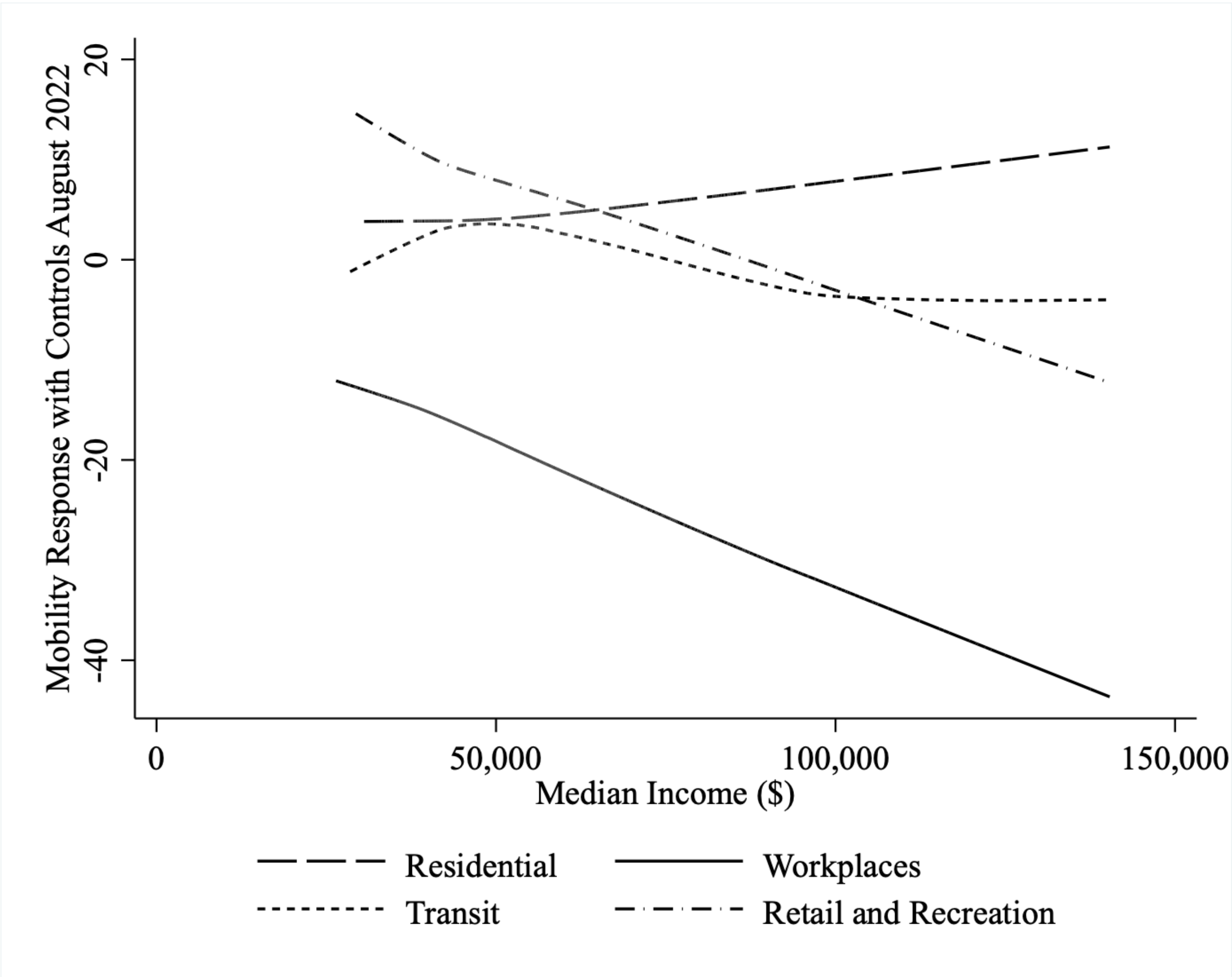
Note: Median household income is annual. Lowess smoothed scatter plots.

Figure 3: Non-parametric sub-functions of median household income in the partial linear regression for the initial phase



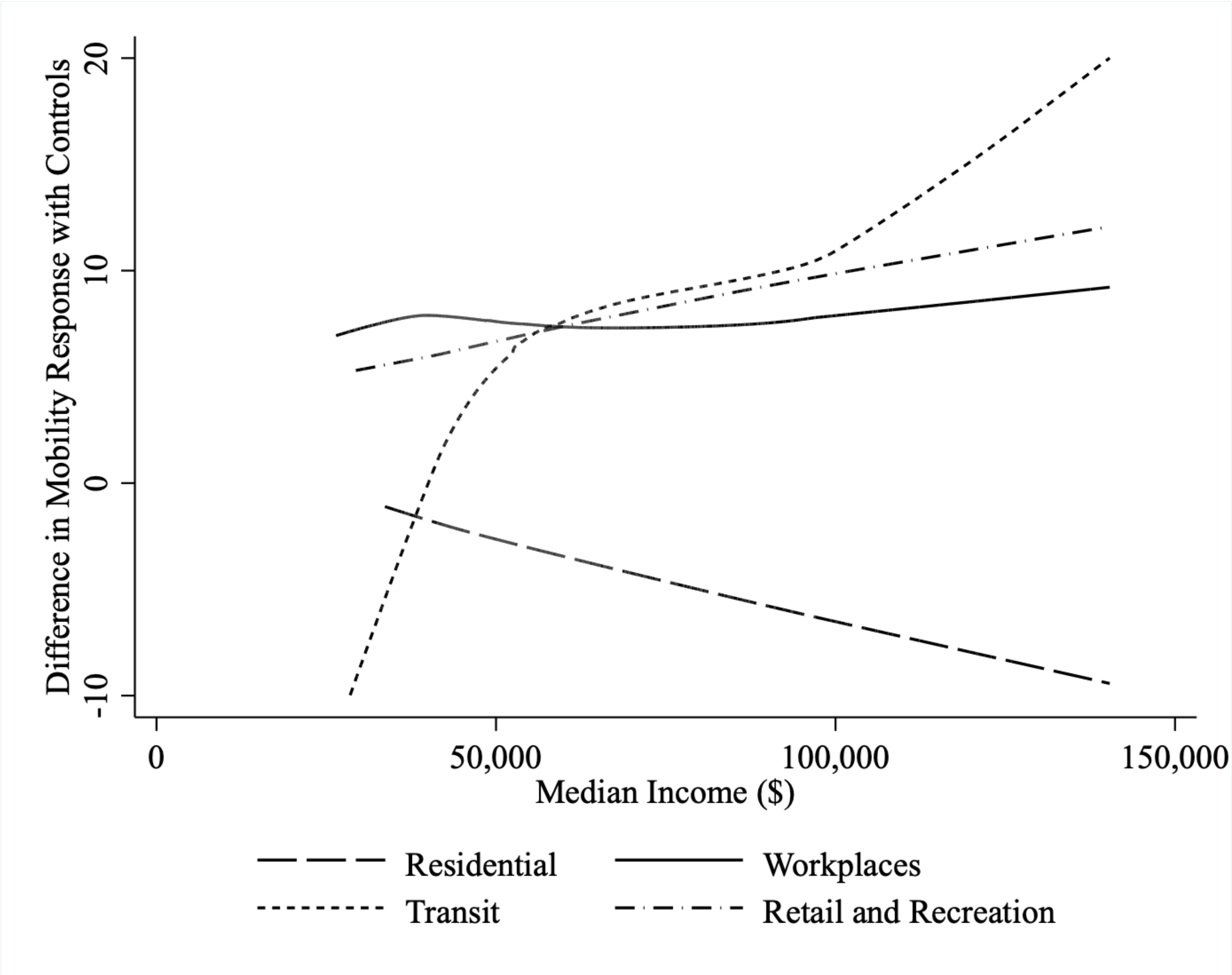
Note: The vertical axis is the estimated non-parametric sub-function of median income ($f^k(.)$ in Equation 6) using PLREG with state fixed effects from Table 2.

Figure 4: Non-parametric sub-functions of median household income in the partial linear regression for the full period up to August 2022



Note: The vertical axis is the estimated non-parametric sub-function of median income ($f^k(.)$ in Equation 6) using PLREG with state fixed effects from Table 4 and estimated on the full period up to 8th August 2022.

Figure 5: Rebound in mobility responses comparing full period with initial phase



Note: The vertical axis gives the estimated non-parametric sub-functions using the change in mobility (normalized by baseline mobility) between the end of the first phase and the endline as the dependent variable, using PLREG with state fixed effects; the coefficients on the linear controls are in Table (5).

Table 1: Summary statistics

	N	Mean	Std. dev.	Min	Max
Mobility: percentage change from the pre-pandemic baseline					
Until June 2020					
Residential	1,380	7.936	3.840	0	28
Workplaces	2,736	-26.682	7.717	-70	-4
Transit stations	988	-3.617	25.298	-90	103
Retail and recreation	1,633	-0.383	15.878	-73	130
Until August 2022					
Residential	1,743	4.483	1.996	-4	15
Workplaces	2,696	-19.260	10.681	-62	66
Transit stations	950	2.774	37.007	-75	230
Retail and recreation	1,549	6.507	20.842	-46	273
Covariates					
Days	3,143	72.53	22.94	0.00	143.00
Days, log	3,143	4.10	0.97	0.00	4.97
Population	3,142	104,127	333,486	88	10,100,000
Population, log	3,142	10.27	1.49	4.48	16.13
Population density	3,143	670.97	4465.49	0.11	179922.30
Potential interaction density, log	3,142	15.00	3.14	3.67	26.40
Population share 65+	3,142	19.27	4.71	4.83	57.59
Population share 65+, log	3,142	2.93	0.25	1.57	4.05
Median income	3,141	52,794.41	13,880.12	25,385.00	140,382.00
Median income, log	3,141	10.84	0.24	10.14	11.85
Poverty rate	3,141	15.16	6.13	2.60	54.00
Poverty rate, log	3,141	2.64	0.40	0.96	3.99
Gini index	3,128	44.55	3.65	25.67	66.47
Gini index, log	3,128	-0.81	0.08	-1.36	-0.41
Share of Black Americans	3,142	9.34	14.47	0.00	86.07
Share of Black Americans, log	3,142	1.61	1.14	0.00	4.47
Share voted for Trump 2016	3,114	0.63	0.16	0.04	0.96
Share voted for Trump 2016, log	3,114	-0.50	0.32	-3.20	-0.04

Notes:
Mobility: Data come from Google's COVID-19 Mobility Reports up to i) 18th of June 2020 and ii) 8th of August 2022. The data measure the change in movement and time spent at each location type relative to the baseline, which is the median value for the corresponding day during a 5-week period beginning in early January. A positive value implies that movement has increased in the places falling under each category relative to the baseline. Retail and recreation include places like restaurants, cafes, shopping centres, theatres, libraries and museums; transit stations include public transportation hubs such as subway, bus, and train stations. N is the number of counties for which we have observations.
Covariates: Days are the number of days since the first case confirmed. Demographic variables are drawn primarily from the U.S. Census Bureau and the CDC. County population density is from this (public) GitHub page. Share of Black American refers to the share of the county population that identify as Black or African American only. Potential interaction density is defined as squared population per square kilometer. The poverty rate for the U.S. is based on the official poverty line. All income variables are pre-pandemic, defined as 2018. Share voted for Trump 2016 refers to the proportion of the county that voted for Donald Trump in the 2016 presidential election.

Table 2: Partial linear regressions for mobility adjustment in response to the pandemic

	Residential		Workplaces		Transit stations		Retail and recreation	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Population interaction density (PID)	0.45***	0.29***	-1.02***	-0.60***	-2.69***	-3.48***	-0.25	-1.11*
	(0.08)	(0.09)	(0.13)	(0.17)	(0.91)	(1.20)	(0.46)	(0.57)
Population	-0.01	0.21	0.31	-0.40	0.41	3.18	-1.44	0.81
	(0.16)	(0.18)	(0.25)	(0.33)	(1.78)	(2.31)	(0.94)	(1.13)
Days since first case	-1.91**	-0.44	0.19	0.39*	8.55***	7.73***	1.00	0.74
	(0.77)	(0.68)	(0.23)	(0.23)	(2.38)	(2.26)	(3.49)	(3.19)
Share 65 and older	-1.81***	-3.01***	6.18***	6.56***	3.78	17.22***	5.24***	14.63***
	(0.32)	(0.32)	(0.56)	(0.59)	(3.90)	(4.54)	(1.97)	(2.08)
Gini index	1.73	2.63**	-3.52*	-4.79***	-50.30***	-39.92***	-22.42***	-35.30***
	(1.23)	(1.12)	(1.85)	(1.82)	(13.84)	(14.10)	(7.51)	(7.26)
Poverty rate	1.36***	1.16***	-4.35***	-3.37***	-7.52	-0.57	-14.15***	-4.72*
	(0.46)	(0.42)	(0.79)	(0.80)	(5.80)	(6.01)	(2.97)	(2.86)
Share of Black Americans	0.20**	-0.36***	0.95***	1.01***	1.10	2.81**	-0.98**	-0.28
	(0.08)	(0.09)	(0.12)	(0.16)	(0.93)	(1.18)	(0.49)	(0.59)
Share voted for Donald Trump 2016	-3.13***	-3.47***	6.79***	7.63***	20.62***	22.89***	10.94***	8.13***
	(0.24)	(0.26)	(0.41)	(0.49)	(2.77)	(3.12)	(1.51)	(1.71)
State fixed effects	No	Yes	No	Yes	No	Yes	No	Yes
N	1368	1368	2715	2715	980	980	1620	1620
f(median income) test statistic	13.18***	13.14***	7.87***	7.65***	0.21	1.48*	2.30**	2.63***

Notes: PLREG with median household income as the continuous variable in the non-parametric sub-function. Data for U.S. counties. The dependent variable comes from Google's COVID-19 Mobility Reports and is of 18th of June 2020. Given that the data are changes from the baseline, a larger negative number indicates more social distancing. All covariates are logged. Robust standard errors in parentheses. See notes to Table 1 for variable descriptions.

Table 3: Regressions for mobility responses ignoring household incomes and their distribution

	Residential		Workplaces		Transit stations		Retail and recreation	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Population interaction density	1.00***	0.85***	-1.47***	-1.45***	-3.02***	-4.13***	-0.40	-1.12**
	(0.08)	(0.10)	(0.11)	(0.16)	(0.73)	(1.02)	(0.38)	(0.49)
Population	-0.45***	-0.23	0.29	0.29	-1.00	1.95	-2.10***	0.13
	(0.18)	(0.20)	(0.23)	(0.31)	(1.46)	(2.00)	(0.79)	(0.98)
Days since first case	2.11**	3.24***	0.01	0.06	9.94***	9.27***	1.00	0.68
	(0.83)	(0.74)	(0.21)	(0.21)	(1.98)	(1.93)	(2.93)	(2.72)
Share 65 and older	-3.70***	-4.82***	8.82***	9.06***	7.10**	13.92***	10.59***	15.96***
	(0.31)	(0.31)	(0.46)	(0.49)	(2.88)	(3.37)	(1.49)	(1.57)
Share of Black Americans	-0.30***	-0.96***	1.32***	1.47***	1.41**	3.78***	-1.67***	-0.28
	(0.07)	(0.09)	(0.09)	(0.14)	(0.69)	(0.98)	(0.36)	(0.49)
Share Donald Trump	-3.48***	-3.19***	8.49***	7.93***	24.96***	24.30***	15.51***	13.00***
	(0.23)	(0.26)	(0.34)	(0.41)	(1.94)	(2.34)	(1.12)	(1.32)
Constant	-4.78	-4.72	-30.16***	-31.23***	10.45	-25.79	7.88	-23.94*
	(3.43)	(3.24)	(2.05)	(2.42)	(13.29)	(16.80)	(12.50)	(12.44)
State fixed effects	No	Yes	No	Yes	No	Yes	No	Yes
R^2	0.612	0.737	0.595	0.646	0.409	0.501	0.334	0.486
R^2 incl. incomes	0.782	0.855	0.681	0.720	0.440	0.518	0.386	0.516
N	1377	1377	2725	2725	986	986	1629	1629

Notes: OLS; robust standard errors in parentheses. All covariates are logged. The second R^2 includes median incomes, the Gini index and the poverty rate.

Table 4: PLREG estimates for the coefficients on the covariates using the full period up to August 2022

	Residential		Workplaces		Transit stations		Retail and recreation	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Density	0.63***	0.15**	-0.36	-0.26	-4.41***	-5.66***	-1.98***	-3.71***
	(0.06)	(0.07)	(0.24)	(0.32)	(1.55)	(2.08)	(0.66)	(0.86)
Population	-0.70***	0.01	-1.35***	-1.63***	1.34	2.98	-1.52	2.13
	(0.12)	(0.14)	(0.48)	(0.63)	(3.00)	(4.01)	(1.36)	(1.71)
Days since first case	-0.93**	-0.20	1.14**	1.29***	3.13	4.53	4.53	6.91
	(0.41)	(0.35)	(0.46)	(0.45)	(4.84)	(4.69)	(5.76)	(5.60)
Share 65 and older	0.37	-0.37	6.05***	7.34***	4.39	6.49	16.24***	16.80***
	(0.25)	(0.25)	(1.04)	(1.12)	(6.70)	(7.89)	(2.86)	(3.21)
Gini index	-0.75	0.10	-14.16***	-12.83***	-36.69	-19.32	-17.97*	-18.31*
	(0.94)	(0.86)	(3.51)	(3.51)	(23.62)	(24.46)	(10.67)	(10.88)
Poverty rate	1.62***	0.99***	-5.38***	-3.07**	-5.30	-1.97	-15.15***	-6.81
	(0.37)	(0.33)	(1.48)	(1.51)	(9.85)	(10.39)	(4.20)	(4.30)
Share of Black Americans	0.13**	-0.14**	0.89***	0.74**	-0.68	0.89	-3.12***	-2.31***
	(0.06)	(0.07)	(0.23)	(0.31)	(1.61)	(2.05)	(0.71)	(0.89)
Share Donald Trump	0.10	-0.76***	5.64***	5.65***	17.37***	17.13***	-3.91*	-0.53
	(0.19)	(0.21)	(0.78)	(0.93)	(4.78)	(5.42)	(2.09)	(2.52)
State fixed effects	No	Yes	No	Yes	No	Yes	No	Yes
N	1730	1730	2677	2677	943	943	1538	1538
F	487.04	130.5	578.23	115.73	76.59	16.80	101.71	28.30

Notes: PLREG with median household income as the continuous variable in the non-parametric sub-function. Data for U.S. counties. The dependent variable comes from Google's COVID-19 Mobility Reports. Given that the data are changes from the baseline, a larger negative number indicates more social distancing. Robust standard errors in parentheses.

Table 5: PLREG estimates of the coefficients on the covariates using the change in mobility (mobility 2022-mobility 2020, normalized by baseline mobility) as the dependent variable

	Residential		Workplaces		Transit stations		Retail and recreation	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Density	0.12	-0.09	0.63***	0.33	-1.01	-1.04	-1.79***	-2.36***
	(0.08)	(0.09)	(0.22)	(0.31)	(1.57)	(2.11)	(0.50)	(0.60)
Population	-0.58***	-0.31*	-1.63***	-1.22**	-0.89	-2.46	0.17	0.97
	(0.16)	(0.17)	(0.45)	(0.59)	(3.05)	(4.06)	(1.02)	(1.20)
Days since first case	-0.34	0.35	1.12***	1.12***	1.55	3.56	-0.05	1.84
	(0.79)	(0.65)	(0.43)	(0.43)	(4.93)	(4.79)	(4.34)	(3.95)
Share 65 and older	2.09***	2.63***	-0.30	0.78	7.05	-3.02	12.97***	3.10
	(0.32)	(0.31)	(0.98)	(1.06)	(6.79)	(8.08)	(2.15)	(2.26)
Gini index	-2.48**	-1.94*	-9.80***	-7.53**	5.98	9.55	4.17	15.37**
	(1.25)	(1.06)	(3.28)	(3.31)	(23.97)	(25.01)	(8.01)	(7.66)
Poverty rate	0.34	-0.29	-1.12	0.25	3.27	0.87	0.26	-0.79
	(0.47)	(0.40)	(1.38)	(1.42)	(10.00)	(10.68)	(3.15)	(3.03)
Share of Black Americans	-0.08	0.27***	-0.07	-0.29	-1.78	-2.35	-2.03***	-1.95***
	(0.08)	(0.08)	(0.21)	(0.29)	(1.64)	(2.13)	(0.53)	(0.63)
Share Donald Trump	3.00***	2.73***	-1.19	-1.98**	-3.99	-6.84	-15.18***	-8.81***
	(0.25)	(0.25)	(0.72)	(0.88)	(4.83)	(5.50)	(1.57)	(1.77)
State fixed effects	No	Yes	No	Yes	No	Yes	No	Yes
N	1367	1367	2674	2674	914	914	1528	1528
F	106.78	42.06	7.35	6.79	2.02	3.08	44.98	17.22

Notes: PLREG with median household income as the continuous variable in the non-parametric sub-function. Data for U.S. counties. The dependent variable is the change in mobility between the endline and the end of the initial phase, normalized by pre-pandemic levels. Robust standard errors in parentheses.