Does Responsiveness to Mortality Risk Vary by Age? Evidence from Pandemic Health Outcomes and Movement Patterns

Ryan Jones Hastings

Professor Christopher Timmins, Faculty Advisor
Professor Michelle Connolly, Faculty Advisor

Honors Thesis submitted in partial fulfillment of the requirements for Graduation with Distinction in Economics in Trinity College of Duke University.

Duke University
Durham, North Carolina
2022

Abstract

When choosing whether to visit venues like stores and restaurants during the COVID-19 pandemic, individuals faced trade-offs between movement and mortality risk. This paper analyzes age-specific responsiveness to infection-related mortality risk in the Philadelphia metropolitan area from March through December 2020. First, we develop a theoretical model that characterizes potential sources of heterogeneity in the decisions of individuals choosing how much to move. Next, we use data on the health outcomes of COVID-19 patients to estimate fatality rates for different demographic groups. Finally, we use a panel of cell phone data tracking visits to venues before and during the pandemic along with a revealed preference approach to estimate an empirical model that relates age to movement decisions. Our results suggest that older people's movements are less sensitive to mortality risk. Under weak assumptions, this implies that older people have a lower willingness to pay for marginal reductions in the probability of death. This finding has implications for the cost-benefit analysis of policies that mitigate adverse health outcomes, such as pandemic movement restrictions and pollution remediation, and for the value of statistical life (VSL) literature more broadly.

JEL classification: D81; I12; J17; R2.

Keywords: COVID-19; VSL; revealed preference; movement; mortality risk; age; cell phone data.

Acknowledgements

While working on this paper, I benefited from the support and guidance of many people. I am grateful for the abundant kindness and mentorship that Chris Timmins has given me over the past three years. He has been exceedingly generous with his time and advice, and I am much the better for it. I am also appreciative of Michelle Connolly for her encouragement and suggestions, which enhanced both the quality of this paper and my enjoyment of the process.

Furthermore, I thank Bill Tobin, Peter Arcidiacono, and Kay Jowers for their investment and insight that have made me a better researcher and more thoughtful person during my time at Duke; the peers in my thesis seminar and the Environmental Justice Lab, especially Julia Marshall, Annabel Hu, Wendy Shi, Nehal Jain, and George Rateb, for their help and humor that have made research less solitary and more fun; and my friends and family, especially my parents, Megan Richards, Alex Rizaldi, Jessie Xu, and John Boom, for their interest in, companionship during, and reassurance about my thesis and whatever I do.

Contents

List of Figures

List of Tables

1 Introduction

The health and mortality risks of infection caused dramatic reductions in aggregate levels of movement during the COVID-19 pandemic. Governments implemented movement restrictions policies and stay-at-home orders with an aim to curb the transmission of the virus by encouraging physical and social distancing. However, individual responses to the pandemic varied dramatically. In the United States, this heterogeneity frequently manifested through political controversy around individual-level decisions to social distance, mask, and receive vaccination.

Because of potential exposure to the virus, individuals faced a trade-off between movement and mortality risk. Choosing to go to venues like stores and restaurants meant facing some risk of COVID-19 infection and, if affected severely, death. When analyzing the reasons individuals make different decisions upon facing this trade-off, it is useful to consider sources of variation in both the mortality risk inherent in movement and the value individuals place on movement.

Several factors make mortality risks heterogeneous across locations. These include variation in the local prevalence of COVID-19 cases, the frequency of other individuals' movements, and the probability of virus transmission at any given destination. Mortality risks also vary across individuals. The probability of death conditional on being infected may vary based on an individual's age and history of underlying medical conditions, as well as the accessibility and quality of medical care in their place of residence, among other factors.

The value individuals place on movement can also vary. Movement may serve to fulfill basic needs like acquiring food or receiving health care. Some individuals employed in occupations deemed essential or with a low ability to conduct work remotely might face a high cost of staying home, potentially foregoing wages and employment opportunities. Those with higher incomes might find it easier to pay for other adaptation strategies like grocery delivery that lead to a lower cost of social distancing and allow them to avoid risking exposure. Other movement may be primarily for recreation or enjoyment purposes.

Another factor in this trade-off is that people may differ in the value they place on reductions in mortality risk. Some people might be risk-averse and try to minimize all chance of infection and death, while others may be less responsive to these risks. This value may differ based on individuals' remaining life expectancies, their expected future earnings, and others' dependence on them for financial or familial support, among other considerations. Note that many of these components vary by age.

It is possible to tease apart the value of reduced mortality risk from occupation- and income-based determinants of the value of movement by looking at non-essential movement decisions, defined as visits to locations such as restaurants or movie theaters that are unlikely to be for employment, nutritional, educational, or medical reasons. Considering only non-essential visits when studying people's risk trade-offs mitigates the concern that heterogeneous movement decisions might be based on people's differing abilities to use adaptation strategies, like remote work, telehealth, and grocery delivery.

In this paper, we leverage heterogeneity in movement patterns and variation in the probability of infection and death from COVID-19 to study how the value of reduced mortality risk varies by age. We use a revealed-preference approach by analyzing microdata on movement patterns during the COVID-19 pandemic to compare the risks people of different ages undertook. The data come from a panel of over 45 million cell phones across the United States that tracks the number of visits to more than 6.8 million venues such as stores, restaurants, and offices.

First, we construct a theoretical model to formulate the trade-off between mortality risk and movement that individuals face. We identify several potential explanations for the heterogeneity in responses seen throughout the pandemic. The model also introduces a mathematically precise definition of the value of reduced mortality risk that motivates the strategy used in the empirical portion of the paper.

Then, we perform an empirical analysis, aiming to quantify how the value of reduced mortality risk varies across age groups. We use data compiled by the Centers for Disease

Control and Prevention (CDC) on the outcomes of patients infected with COVID-19 to estimate the fatality rate of the virus for different demographic groups. We also consider the role that geographic characteristics, like the availability and quality of medical care, might play in the probability of death conditional on infection. We couple this analysis with data on spatial and temporal variation in the prevalence of the virus to construct an empirical measure of mortality risk.

Then, equipped with this measure, we examine patterns in distancing as tracked by cell phone movement data. We classify venues based on their industries, as defined by their associated North American Industry Classification System (NAICS) codes, to differentiate between essential and non-essential visits. We use block group-level demographic data from the 2015-2019 American Community Survey to learn about the age profiles of the block groups from which observed cell phones originate. Then, we perform a regression analysis of non-essential movement frequency to estimate age-varying effects of mortality risk on movement. Our findings suggest that older people are less responsive to mortality risk. Using comparative statics, we show that under weak assumptions, this implies that older people have a lower willingness to pay for marginal reductions in the probability of death.

The value of marginal reductions in the probability of death, frequently studied by economists in the literature on the value of statistical life (VSL), is an important parameter in the cost-benefit analysis of policies and regulations involving any sort of health or fatality risk. Estimates of this value are frequently used by federal agencies. For example, the Environmental Protection Agency weighs the health benefits of environmental policies like pollution remediation, the Department of Transportation considers the fatality risks associated with traffic hazards and infrastructure quality, and the Department of Health and Human Services reviews policies like movement restrictions that aim to mitigate adverse health outcomes. This paper builds on previous VSL studies by using a novel setting and approach to characterize heterogeneity in the value of reduced mortality risk.

2 Literature Review

This paper contributes to two distinct literatures. First, it contributes to understanding about age-based heterogeneity in the value of reduced mortality risk. Second, it adds to a rapidly growing body of work examining individual mobility responses to the COVID-19 pandemic.

2.1 Value of Statistical Life

Patterns in the value of reduced mortality risk have been studied extensively in the economic literature on the Value of Statistical Life (VSL). VSL is defined as the marginal rate of substitution between mortality risk and wealth. To study this concept, economists typically observe the risk-money trade-off for small reductions in the probability of death. It measures an individual's willingness to pay for a marginal reduction in mortality risk. VSL is a revealed-preference non-market valuation technique entirely separate from the calculation of legal compensation for injury or death. VSL is often used in cost-benefit analysis to place monetary values on the mortality effects of policy proposals. Economists have inferred VSL estimates from a variety of contexts including trade-offs between wages and workplace fatality risks, house prices and air pollution health risks, and travel time savings and automobile accident rates; for reviews of VSL techniques and estimates, see Hammitt (2000), Viscusi and Aldy (2003), Blomquist (2004), Ashenfelter (2006), and Viscusi (2012). These estimates can then be used to evaluate the costs or benefits of, for example, workplace safety regulations, pollution remediation, and traffic policies.

There is no pecuniary exchange observed in the context of decisions to move or social distance during the pandemic, so we do not construct a monetary estimate. However, we still observe changes in risky behavior, in this case movement, in response to changes in fatality risk, in this case the probability of infection and death from COVID-19. These trade-offs can give us information similar to those used in the VSL literature and allow us to examine the relative magnitudes of the value of mortality risk reductions for different age groups, as revealed by their movement decisions. Throughout the paper, we use the term Value of Reduced Mortality Risk (VRMR) instead of VSL because of criticisms of the latter's clarity and ethical implications.[1]

Several papers have studied how estimates of VRMR vary by age. Most are summarized in Aldy and Viscusi (2007). The leading hypothesis is that the distribution of VRMR conditional on age follows an inverse-U shape (Ehrlich and Yin, 2005; Kniesner et al., 2006; Aldy and Viscusi, 2008). That is, VRMR tends to grow, peak, and eventually decline throughout the life cycle. However, some models have shown the pattern by age to be theoretically ambiguous (Johansson, 1996, 2002) and one study found that higher compensation was needed for older workers to accept small increases fatality risks, contrary to the inverse-U hypothesis (Smith et al., 2004). The intuition behind the eventual decline is that older individuals are essentially purchasing fewer additional years of life expectancy. The decisions to avoid mortality risk can be thought of as investments in future time and consumption. As such, intuition for the inverse-U hypothesis can come from the notion that age patterns in VRMR may track the life-cycle pattern of consumption (Kniesner et al., 2006). As consumption tends to increase with age, people may be more willing to pay for a reduction in mortality risk as they get older. Eventually, consumption may flatten out and the effects of lower remaining life expectancy may dominate. We aim to test the inverse-U hypothesis by using pandemic movement patterns to estimate (relative) VRMR across age groups.

These age patterns have significant policy implications. For example, in 2003, the US Environmental Protection Agency began using a "senior discount" when assessing the benefits of pollution remediation policies because older people had lower estimated VSLs (Viscusi,

[1]The use of VSL, especially in cost-benefit analysis, has faced extensive critique. Some scholars invoke these ethical issues inherent in trying to put monetary values on people's lives (Broome, 1978; Heinzerling, 2000). There are counter-arguments to the points of these critiques (Viscusi, 2009b), but we largely avoid this issue by not attempting to estimate monetary values. A more salient critique is that its name can be off-putting or misleading (Cameron, 2010). Simon et al. (2019) organize focus groups to determine the best alternative term for the same concept; they find value of reduced mortality risk (VRMR) to be the preferred term because it more effectively communicates the concept's meaning. As a a result, we use the term VRMR instead of VSL throughout the paper.

2009a). One paper has argued that a similar discount should be applied when assessing the benefits of COVID-19-related policies because the mortality effects being considered disproportionately affect older individuals who are believed to value mortality risk reductions less (Hammitt, 2020).

2.2 COVID-19 and Mobility

There is a burgeoning literature studying mobility responses to the COVID-19 pandemic, including papers both theoretical and empirical in nature.

A number of papers construct theoretical models of decisions to social distance during a pandemic. Allcott et al. (2020a) formulate the decision as a dynamic optimization problem and identify three dimensions that could cause heterogeneity in movement patterns: the marginal utility of movement, the marginal infection probability, and the private cost of infection. Several other papers model decisions to social distance using game theory to allow for social interactions and externalities (Vandenbroucke, 2020; Gosak et al., 2021; Reluga, 2010; Alfaro et al., 2020). We use aspects of these papers' approaches when constructing our model highlighting potential determinants of movement decisions.

The empirical literature on movement has three main focuses. First, it examines transmission patterns of the virus. Second, it examines the effects of policy interventions like movement restrictions. Third, it examines heterogeneity in movement patterns based on demographics and other geography-specific characteristics. Our contribution is to this third sub-literature. Like in this paper, much of this literature uses cell phone data to study movement patterns.

Several papers study the effects of policy responses–including stay-at-home orders, lockdowns, emergency declarations, and reopening–on movement patterns (Glaeser et al., 2021; Yechezkel et al., 2021; Gao et al., 2020; Alexander and Karger, 2020; Goolsbee and Syverson, 2020; Allcott et al., 2020b; Engle et al., 2020). While most find that more restrictive policies decrease movement patterns, Goolsbee and Syverson (2020) find that individual choices are a more important determinant of movement decisions than legal orders. These effects are difficult to disentangle because individual and policy responses evolved simultaneously. This motivates the need to examine determinants of individual decisions more closely.

Studies examining the heterogeneity of movement patterns across demographics have focused on a number of different determinants, including income (Weill et al., 2020; Chang et al., 2021; Lamb et al., 2021; Jay et al., 2020; Chiou and Tucker, 2020; Kavanagh et al., 2021), race (Chang et al., 2021; Engle et al., 2020; Couture et al., 2021), occupation (Mongey et al., 2020), partisanship (Allcott et al., 2020a; Barrios and Hochberg, 2020; Kavanagh et al., 2021; Engle et al., 2020), and internet access (Chiou and Tucker, 2020). However, there is limited research examining associations between movement patterns and age. We aim to analyze this relationship with more nuance.

3 Theoretical Framework

In this section, we consider potential sources of heterogeneity in individuals' decisions when facing the trade-off between movement and mortality risk. Then, we use the utility-theoretic definition of the value of reduced mortality risk (VRMR) to show how we can learn about VRMR by observing these trade-offs.

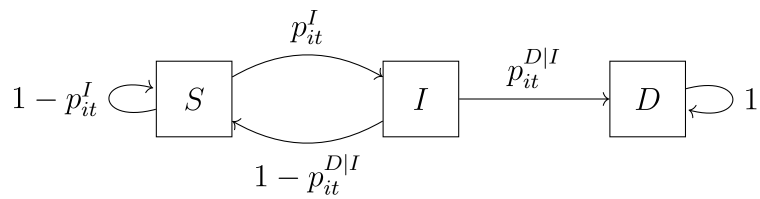
3.1 Determinants of Mortality Risk

First, we analyze determinants of mortality risk. We consider both the amount of exposure to COVID-19 during movement and the likelihood of infection and death conditional on exposure.

S-I-D Model

We use a simple Susceptible-Infected-Dead (S-I-D) epidemiological model to represent the spread of COVID-19, as depicted in Figure 1. We allow infection and mortality risk to vary across individuals $i \in \mathcal{I}$ and weeks $t \in \mathcal{T}$. A susceptible individual becomes infected with probability of infection p^I_{it}, which can vary across individuals and time periods, and otherwise remains susceptible. An infected indivdual dies with probability of death conditional on infection $p^{D|I}_{it}$, which can also vary across individuals and time periods, and otherwise returns to being susceptible. Death is an absorbing state; individuals remain with probability 1.

Figure 1: S-I-D Model of COVID-19 Spread



Notes: This directed graph highlights our simple epidemiological model of COVID-19 spread. Individuals can be in one of three states: susceptible, infected, or dead. Susceptible individuals become infected with probability p_{it}^{I}, which can vary across individuals and time periods. Infected individuals die with probability $p_{it}^{D|I}$, which can also vary across individuals and time periods.

Since COVID-19 infection lasts roughly one to two weeks, we use t to index weeks and do not allow individuals to remain infected for multiple time periods. We make the simplifying assumption that all individuals who have previously been infected remain susceptible, as opposed to being recovered. Despite some evidence of immunity following infection, many individuals contract COVID-19 multiple times, especially as different variants arise.

Mortality risk, p_{it}^{D}, is defined as the probability that a susceptible individual dies and is equal to the product of the probability of infection and the probability of death conditional on infection, as show in equation 1. We will examine determinants of each component and impose additional structure on the heterogeneity across individuals.

$$p_{it}^{D} = p_{it}^{I} \cdot p_{it}^{D|I} \tag{1}$$

Sources of Heterogeneity

Exposure. Individuals can be exposed to COVID-19 when visiting venues like stores and restaurants. The amount of exposure at venue $k \in \mathcal{K}$ depends on three factors: $s_{j_k t}^{I}$, the share of individuals in geography j_k (in which k is located) that are infected with COVID-19 a time t; n_{kt}^{V}, the number of other visitors to venue k; and c_k, a site-specific factor measuring the amount of contact among visitors, which may depend on the physical characteristics of the space such as square footage and whether it is indoors or outdoors. Venue-specific exposure is represented by equation 2.

$$e_{kt} = s^I_{j_k t} \cdot n^V_{kt} \cdot c_k \tag{2}$$

In a given period, individuals may visit multiple venues. $\mathcal{K}_{it} \in \mathcal{P}(\mathcal{K})$, the set of venues individual i visits at time t, is an element of the power set of $\mathcal{K}$. The total exposure of an individual y_{it} is given by the sum of the relevant venue-specific exposures, as shown in equation 3.

$$y_{it} = \sum_{k \in \mathcal{K}_{it}} e_{kt} \tag{3}$$

Probability of Infection. Per unit of exposure y_{it}, infection occurs at a rate $r_t(x_i)$ that we allow to depend on individual characteristics x_i and time t. The infection rate may vary among individuals based on demographic and health factors, including age, race, sex, general health status, and underlying medical conditions. It is also affected by any adaptation strategies, such as masking, used by individuals to mitigate the chance of infection when exposed to an infected individual. The infection rate may vary across time periods because of changes in the contagiousness of the virus and the mitigation efforts used to prevent its spread. The probability an individual i is infected at time t is given by equation 4.

$$p^I_{it}(y_{it}) = y_{it} \cdot r_t(x_i) \tag{4}$$

Probability of Death Conditional on Infection. Similarly, the fatality rate $f_t(x_i, z_j)$ can depend on individual characteristics x_i and geographic characteristics z_j. Individual demographics–including race, age, and sex–and health factors–including general health status, physical inactivity, substance use, and underlying conditions–may affect fatality rate. Access to and quality of medical care also determines fatality rate. This varies across geographic areas based on the spatial distribution of care providers and care quality. The accessibility of medical care also varies among individuals based on characteristics like income, insurance coverage, internet access, vehicle ownership, and language ability. The fatality rate could also vary over time as variants of the virus that have mutations affecting its virulence develop or as local medical facilities operate below, near, or at capacity. However, it does not depend on the amount of movement y_{it}. The probability of death conditional on infection is given by equation 5.

$$p_{it}^{D|I} = f_t(x_i, z_j) \tag{5}$$

Summary. Combining equations 1 through 5, mortality risk is given by equation 6. Individuals' mortality risks depend on the set of venues they choose to visit $\mathcal{K}_{it}$, those venues' characteristics c_k, the number of others visiting the same venues n_{kt}^V, the prevalence of COVID-19 in the area $s_{j_k t}^I$, and the infection and fatality rates $r_t(x_i)$ and $f_t(x_i, z_j)$ that vary based on demographic characteristics, health factors, and geographic characteristics.

$$p_{it}^D = \left(\sum_{k \in \mathcal{K}_{it}} s_{j_k t}^I \cdot n_{kt}^V \cdot c_k \right) \cdot r_t(x_i) \cdot f_t(x_i, z_j) \tag{6}$$

A Mental Model: Simplifications for Estimation

We make a number of simplifying assumptions. These simplifications serve two purposes: to better capture how individuals may *perceive* mortality risk and to facilitate the connection between the theoretical model and the data we have for our empirical work.

Exposure. Although we discussed how exposure may vary across venues based on the number of other visitors and location-specific characteristics, we make the simplifying assumption that the exposure is identical at all locations: $e_{kt} = 1 \ \forall k$. Substituting this assumption into equation 3, the total exposure of an individual is equal to the number of venues visited, as shown in equation 7 where the set norm $|\cdot|$ measures the size of the set of venues visited.

$$y_{it} = \sum_{k \in \mathcal{K}_{it}} e_{kt} = \sum_{k \in \mathcal{K}_{it}} 1 = |\mathcal{K}_{it}| \tag{7}$$

Under this assumption, we can now refer to y_{it}, previously defined as the amount of exposure, as the amount of movement. We use the number of visits as the choice variable in an individual's movement decision and the outcome variable in our empirical analysis. By making this assumption, we do not consider how individuals may adjust what venues they visit and what times they choose to go based on the numbers of contemporaneous visitors and location-specific characteristics; instead, we suggest that their main decision is whether and how often to go to venues in the first place.

Infection per Unit Exposure. Now that we have made this assumption about exposure, we also adapt and make assumptions about the infection rate per unit exposure $r_t(x_i)$. Specifically, we suggest that this value–which is now simply the infection rate per visit–is proportional to the share of people infected s^I_{jt} in an individual's place of residence j at time t: $r_t(x_i) \propto s^I_{jt}$. Substituting this assumption and equation 7 into equation 4, we see in equation 8 that the probability of infection is proportional to the product of the number of locations visited and the share of infected residents.

$$p^I_{it}(y_{it}) = y_{it} \cdot r_t(x_i) \propto |\mathcal{K}_{it}| \cdot s^I_{jt} \tag{8}$$

This assumption is likely reflective of how individuals perceive infection risk: they update their expectations of infection based on publicly available information about the prevalence of the virus in their area. We also impose that the key variation in the probability of death across demographic groups derives from the probability of death conditional on infection rather than the probability of infection.

Equation 9 updates equation 6 based on these assumptions. Now, an individual's probability of death is proportional to the number of visits they make $|\mathcal{K}_{it}|$, the share of infected residents s^I_{jt}, and the fatality rate $f_t(x_i, z_j)$, which can depend on individual and geographic characteristics.

$$p^D_{it} \propto |\mathcal{K}_{it}| \cdot s^I_{jt} \cdot f_t(x_i, z_j) \tag{9}$$

3.2 Trade-Offs in Movement Decisions

Next, we characterize the trade-offs between movement and mortality risk by modeling individuals as utility-maximizing agents. We leave out subscripts, but we consider the choice of an individual i in a given time period t. We assume that individuals are myopic, not forward-looking, and so use a static model of movement choice.

Utility-Maximizing Problem

Individuals choose both where and how much to move. In line with the assumptions in the previous section, we focus on their choice of the number of venues visited, y.

Objective Function. For a level of movement y, individuals get some positive utility from moving relative to staying home, $u^M(y, x)$, which may vary based on individual characteristics x. They receive disutility $d^I(x)$ from being infected, which occurs with probability $p^I(y, x)$, and receive disutility $d^D(x)$ from dying, which occurs with probability $p^D(y, x)$. These disutilities may also vary based on individual characteristics.

Individuals choose the amount of movement to maximize utility from moving less the potential disutility from infection and death. An individual's utility maximizing problem is given by equation 10. We assume that individuals generally perceive the infection risk per visit correctly, since there is publicly available data on new COVID-19 cases and mortality risks. We also assume that individuals perceive fatality rates correctly; we talk about potential implications of this assumption in our discussion section. Substituting equation 1 into equation 10 gives equation 11, and substituting equations 4 and 5 into equation 11 gives equation 12, where $r(x)$ is the marginal increase in the probability of infection per unit of y and $f(x,z)$ is the fatality rate conditional on being infected.

$$\max_{y} U = u^M(y,x) - p^I(y,x)d^I(x) - p^D(y,x)d^D(x) \tag{10}$$

$$= u^M(y,x) - p^I(y,x)\left[d^I(x) + p^{D|I}(x)d^D(x)\right] \tag{11}$$

$$= u^M(y,x) - y \cdot r(x)\left[d^I(x) + f(x,z)d^D(x)\right] \tag{12}$$

First-Order Condition. The first-order condition of the objective function with respect to y reveals that utility-maximizing agents will choose the amount of movement y such that equivalent equations 13 and 14 hold. Individuals equate the marginal utility of moving with the expected marginal disutility of infection and death.

$$\underbrace{\frac{\partial u^M}{\partial y}(x)}_{\text{marginal utility of moving}} = \underbrace{\frac{\partial p^I}{\partial y}(x)}_{\text{marginal infection rate}} \cdot \underbrace{\left[d^I(x) + p^{D|I}(x) \cdot d^D(x)\right]}_{\text{expected disutility of infection}} \tag{13}$$

$$= r(x) \cdot \left[d^I(x) + f(x,z) \cdot d^D(x)\right] \tag{14}$$

Divergent movement responses across age groups could be because of any of these three factors. We have previously discussed individual characteristics that could affect the marginal infection rate $r(x)$ and the fatality rate $f(x,z)$. We now discuss how the utility of moving and expected disutility of infection might vary among individuals.

Sources of Heterogeneity

Utility of Moving. We specify three key determinants of the utility of moving u^M, although there may be others. First, movement can serve to fulfill basic needs. These include food, healthcare, and education. Second, movement may be for employment purposes. Someone may have to move to get paid or maintain their employment status. Third, movement can have recreation value. This captures a person's enjoyment of and entertainment from the activity.

Note, however, that the contributions of the first two components depend on the availability and use of adaptation strategies. Food delivery, telehealth, remote learning, and remote work all decrease the utility of movement relative to staying home since an individual can meet basic needs or continue employment without movement. However, not all may have the choice to use these.

In this paper, we focus on non-essential visits in order to avoid considerations of movement for basic needs and employment purposes, since the availability of adaptation strategies is a function of a person's geographic location, socio-economic status, and demographic characteristics. Therefore, the utility from a non-essential visit depends only on the recreation value of the visit relative to staying at home and the expected disutility from potential adverse health outcomes.

Although the consideration of only non-essential visits alleviates some concerns, there are two that remain when trying to analyze how people differentially value reductions in mortality risk based based on age. First, the value of the outside option (staying at home) may vary systematically across age groups. The value of staying home could depend on a variety of characteristics such as the size and quality of housing, the number of other people living and working in that space, and the availability of proximate parks and outdoor space. This would be an issue when examining age-based relative movement patterns if, for example, older individuals tend to have higher quality houses or are more likely to be in homes and communities (like assisted living facilities) that have high amenity provision. Second, individuals' enjoyment of non-essential activities may systematically differ across age groups. For example, if we were considering restaurants as the universe of potential locations individuals could visit, it would be fine if people of different ages have different tastes *among* restaurants (they might prefer some to others) but would be an issue if people of different ages have different tastes *for* restaurants (one age group tends to like restaurants more and others tend to like them less).

Expected Disutility of Infection. The disutility of infection d^I reflects the psychological and financial costs of infection. The financial costs could include both direct medical costs and opportunity costs. For example, there could be a higher cost of infection for individuals who would lose their job or not get paid during the time that they are sick.

The disutility of death d^D reflects how generally averse someone is to dying. This value could be based on the financial costs of death, in terms of lifetime consumption lost. We would expect the financial cost of death to be lower for older people because they have fewer years of life expectancy. It could also depend on the effects of death on any kids or dependents; people with young children might have greater disutility from dying because it would have a negative impact on their kids. Finally, it could reflect how much an individual prioritizes current quality of life. Older people may care more about having a high quality of life now and have a high discount factor for future utility. The more an individual discounts future utility, the smaller the effect of death on expected lifetime utility would be.

3.3 Value of Reduced Mortality Risk

Previously, we characterized an individual's trade-off between movement and mortality risk and potential sources of heterogeneity. Now, we describe how this trade-off relates to the value of reduced mortality risk (VRMR), a key parameter used in the cost-benefit analysis of policies mitigating adverse health outcomes. An individual's VRMR, which represents their willingness to pay for a marginal reduction in mortality risk, is defined as the marginal rate of substitution between mortality risk p^D and wealth w, as seen in equation 15.

$$VRMR \equiv \frac{dw}{dp^D} = \frac{\partial U/\partial p^D}{\partial U/\partial w} \tag{15}$$

When analyzing movement, we do not observe a pecuniary trade-off, so we cannot directly evaluate the relationship between wealth and mortality risk. Instead, we use a revealed preference approach to learn about the marginal rate of substitution (MRS) between mortality risk and movement, dy/dp^D, defined in equation 16.

$$\frac{dy}{dp^D} = \frac{\partial U/\partial p^D}{\partial U/\partial y} \tag{16}$$

Equation 17 shows the relationship between VRMR and the MRS of mortality risk and movement. There is an additional term affecting VRMR: the marginal rate of substitution between movement and wealth, which is the marginal willingness to pay for an additional unit of movement. This equation reveals that to conclude anything about the way VRMR varies across ages using observed trade-offs between movement and mortality risk, we must make assumptions about the marginal rate of substitution between movement and wealth.

$$VRMR = \frac{\partial U/\partial p^D}{\partial U/\partial w} = \frac{\partial U/\partial p^D}{\partial U/\partial y} \cdot \frac{\partial U/\partial y}{\partial U/\partial w} = \underbrace{\frac{dy}{dp^D}}_{\text{revealed}} \cdot \underbrace{MWTP_y}_{\text{assumed}} \tag{17}$$

Later in the paper, after estimating the first component, we perform comparative statics to see what we can conclude about VRMR under different assumptions of the way the marginal willingness to pay for movement varies by age.

We have highlighted several potential sources of heterogeneity in movement decisions, including variation in mortality risk, the utility of moving, and the disutilities of infection and death. To parse these out to some extent, we quantify mortality risk and make assumptions about the utility of moving using shape restrictions on the marginal willingness to pay for movement such that we can learn about people's willingness to pay for marginal reductions in the probability of death. However, previous studies that have examined how VRMR varies by age have been agnostic to the exact mechanisms behind the patterns they find. Even without identifying each component separately, it is fruitful to discuss the composite effects.

4 Empirics: Mortality Risk

In this section, we empirically study mortality risk using data on infection rates and patient health outcomes. We observe patient-specific health outcomes and location-specific characteristics, which we use to estimate a logistic regression model of the probability of death conditional on being infected. We also observe time-varying county-level COVID-19 case rates, which we combine with the logit estimates to construct a measure of mortality risk for an individual in a given location at a given time, which differs across age and other demographic groups.

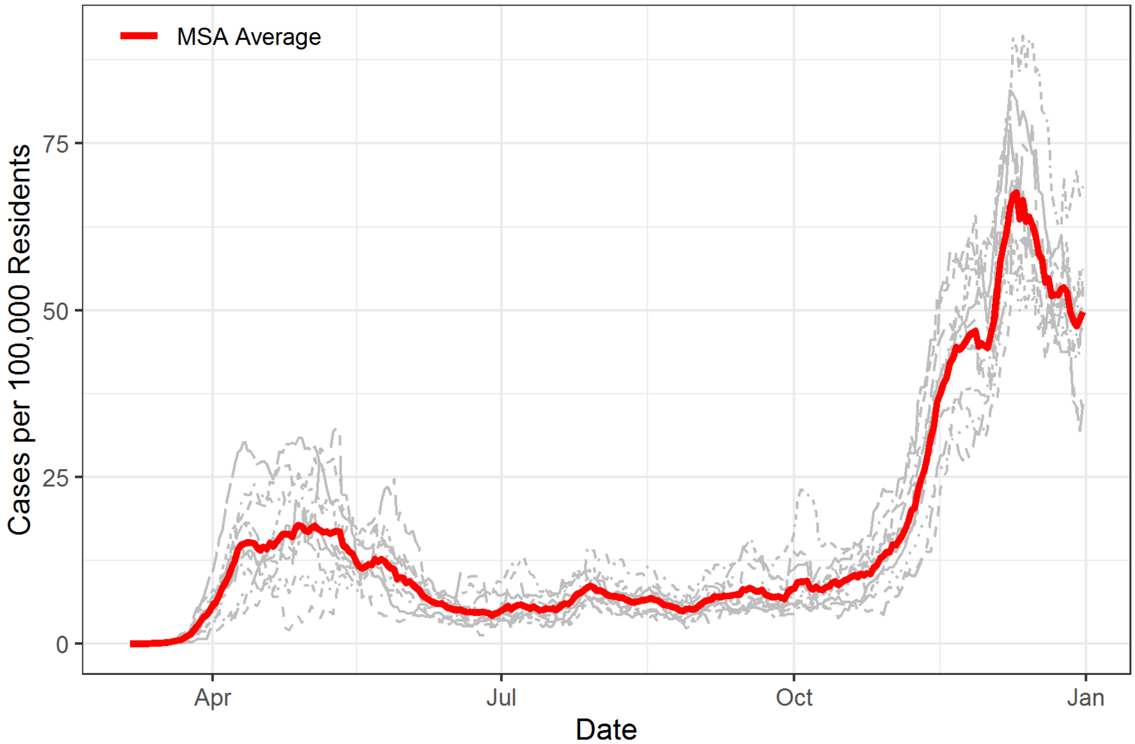
4.1 Data

We describe our data on the determinants of mortality risk, including infection and fatality rates. We consider heterogeneity across individuals and geographies.

Infection Rates. We measure COVID-19 prevalence using county-level data compiled by the *New York Times* from state and local health agencies. The data report seven-day rolling averages of new COVID-19 cases and deaths for January 21, 2020 to the present.[2] Infection rates will be a component of our constructed mortality risk measure that we use in our empirical study of movement decisions, so it is important to note the sources of variation in this variable. Figure 2 plots the evolution of case rates in the Philadelphia Metropolitan Statistical Area (MSA) from March to December 2020, displaying the variation within and across time perods. We see that there is substantial cross-sectional variation, especially in the months of April, May, November, and December. There is also variation over time, with two peaks in April-May and November-December. However, there is a significant stretch of time between July and October where the average MSA case rate hovers between six and eight cases per 100,000 residents.

[2] These data can be accessed at https://github.com/nytimes/covid-19-data.

Figure 2: Time Series of COVID-19 Infection Rates in Philadelphia



Notes: This plot shows the variation in COVID-19 infection rates in the Philadelphia Metropolitan Statistical Area (MSA) within and across time periods from March through December 2020. Each grey line displays the seven-day rolling averages of new cases per 100,000 residents in one of the eleven counties in the MSA. The red line plots the average MSA case rate over time, calculated by taking the mean of the eleven county-specific case rates. The plot is made using the *New York Times* COVID-19 data.

Health Outcomes. We use the COVID-19 Case Surveillance Public Use Data published by the US Centers for Disease Control and Prevention (CDC) for information on the severity and fatality rate of COVID-19 infections. The dataset has patient-level data on millions of COVID-19 cases reported by state and local health departments.

The data are available in two forms.[3] In the first form, which contains no geographic information, we observe each patient's age group,[4] race,[5] gender, the month in which the case

[3]These data are available as the COVID-19 Case Surveillance Public Use Data (without geography), https://data.cdc.gov/Case-Surveillance/COVID-19-Case-Surveillance-Public-Use-Data/vbim-akqf, and the COVID-19 Case Surveillance Public Use Data with Geography, https://data.cdc.gov/Case-Surveillance/COVID-19-Case-Surveillance-Public-Use-Data-with-Ge/n8mc-b4w4.

[4]The age bins reported are 0-9, 10-19, 20-39, 40-49, 50-59, 60-69, 70-79, and 80+.

[5]The race and ethnicity categorizations we construct are: White, Non-Hispanic; Asian, Non-Hispanic;

occurred, and health outcomes including whether the patient was hospitalized or died. In the second form, we also observe the individuals' county and state of residence, but only see a coarser classification of age group.[6] Given that this paper's focus is how people of different ages respond to mortality risk, it is important to get precise estimates of how fatality rate (and thus mortality risk) vary by age, which requires using the finer age bins. However, as a robustness check, we examine the sensitivity of the fatality rate estimates for different age groups to the inclusion of geography-specific factors like access to medical care and the prevalence of related health conditions. We can only perform this analysis with the data that has coarser age bins; therefore, we use both forms of the data.

Geographic Characteristics. Although we observe age, sex, and race in the CDC data, in the theoretical framework we identify a number of other significant determinants of fatality rates, including general health status, underlying conditions, and access to and quality of medical care. We use aggregate characteristics of the county in which a patient lives to determine the influence of these other factors. We compile county-level data from two sources: the County Health Rankings (CHR) Analytic data, compiled by the University of Wisconsin Population Health Institute from a number of government sources, and the American Community Survey (ACS) 2015-2019 5-year estimates. We focus on indicators of the prevalence of populations that the CDC has identified as needing extra precautions because of incread risk for severe COVID-19 illness[7] and variables highlighted in the CDC Social Vulnerability Index,[8] which aims to identify communities that need additional support during disasters (including disease outbreaks) because of heightened health stressors or limited access to medical care.

The CHR variables we consider are: the percent of adults who are current smokers, the percent of the adult population that are obese, the number of primary care physicians per

Black, Non-Hispanic; Native Hawaiian, Pacific Islander, American Indian, or Alaska Native, Non-Hispanic; Multiple/Other, Non-Hispanic; and Hispanic/Latinx.

[6] The age bins reported are 0-17, 18-49, 50-64, and 65+.

[7] https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/index.html

[8] https://www.atsdr.cdc.gov/placeandhealth/svi/index.html

100,000 people, the percent of the population under age 65 without health insurance, and the average daily density of fine particulate matter (PM2.5). The ACS variables we consider are: the median household income, the percent of the population with income below the poverty level, the percent of the population in a rural area, the percent of households who have limited English speaking ability, the percent of households with no vehicle available, and the percent of households without access to a computer with an Internet subscription.

4.2 Model

We describe how we use this data to estimate fatality rates for different demographic groups and construct a measure of mortality risk.

Fatality Rates. We use a logistic regression model with the CDC data to estimate heterogeneous fatality rates. Specifically, we parameterize the fatality rate using the functional form of equation 18, which allows the probability of death conditional on infection to vary based on a vector of individual demographics $\boldsymbol{x}_i$–comprised of age, sex, and race–and a vector of geographic characteristics $\boldsymbol{z}_j$.

$$f(\boldsymbol{x}_i, \boldsymbol{z}_j) = logit^{-1}\left(\zeta + \sum_a \eta_a 1_{[Age_i=a]} + \sum_s \eta_s 1_{[Sex_i=s]} + \sum_r \eta_r 1_{[Race_i=r]} + \boldsymbol{\omega} \boldsymbol{z}_j\right) \tag{18}$$

We use logistic regression on observations with non-missing death outcomes to obtain parameter estimates $(\hat{\zeta}, \hat{\boldsymbol{\eta}}, \hat{\boldsymbol{\omega}})$ which we can use to predict fatality rates $\hat{f}(\boldsymbol{x}_i, \boldsymbol{z}_j)$. Values of $\hat{\boldsymbol{\eta}}$ may reflect differences in biological risks, access to medical care, and any other health determinants correlated with age, race, and sex. However, omitted variable bias is not an issue since we only aim to predict fatality rates for use in our analysis of movement decisions and are not trying to isolate any causal effects.

Infection Rates. We approximate the share of infected visitors in block group j at time t with the county-level seven-day rolling average of new cases per 100,000 residents, $\hat{s}^I_{jt}$, from the *New York Times* data.

Mortality Risk. Recall from equation 9 that under our model assumptions, an individual's marginal probability of death for an additional venue visit is proportional to the product of the share of infected residents and the fatality rate. With empirical estimates of fatality rate and infection rate, we can construct an empirical measure of the marginal mortality risk per unit of movement using equation 19. We define this measure as M.

$$\hat{p}^D_{ijt} \propto \hat{s}^I_{jt} \cdot \hat{f}(\boldsymbol{x}_i, \boldsymbol{z}_j) \equiv M \tag{19}$$

Note that the time variation of this measure comes entirely from variation in the share of infected residents. Also note that the share of infected residents is measured at the county level, so the only variation in this measure at geographic units finer than the county level comes from the estimates $\hat{f}$. Appendix A describes the aggregation of this measure M from the individual level to the block group level in more detail.

4.3 Results

Table 1 shows the parameter estimates of the logistic regression model for the probability of death conditional on infection using the CDC data that has fine age bins and no geographic information. The excluded categories, which determine the baseline to which other categories are compared to, are age 20-29, race white, and sex female. We see that the coefficients for age groups are monotonically increasing in age: older age groups all have positive coefficients which increase in magnitude as age increases. This implies that the fatality rate is higher for older individuals. We also see that the coefficient on male is positive, implying that men face a higher fatality rate than women, and that the coefficients on all other races are positive, implying that non-white individuals face higher fatality rates than white individuals. All coefficients are statistically significant at the 0.01 level.

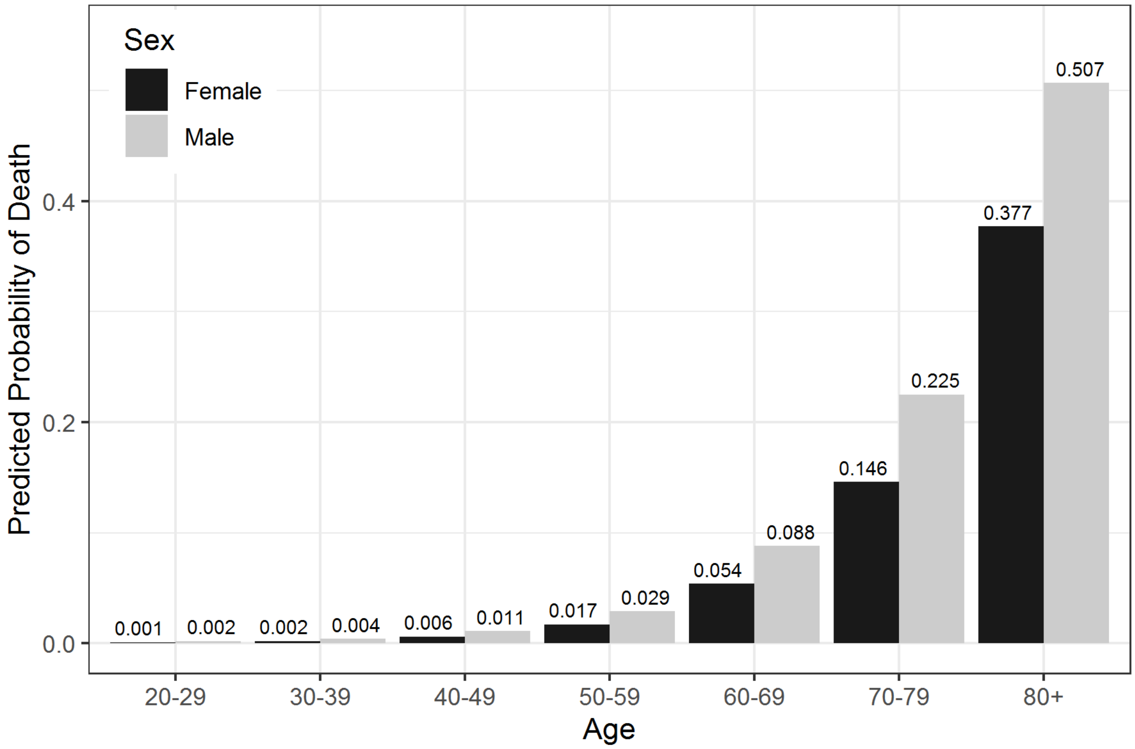
Table 1: Fatality Rate Parameter Estimates: Demographics

Dependent Variable:		Death	
Model:	(1)	(2)	(3)
Variables			
(Intercept)	-6.279***	-6.549***	-7.025***
	(0.0240)	(0.0241)	(0.0243)
Age = 0-9	-0.8376***	-0.8715***	-0.9584***
	(0.0817)	(0.0817)	(0.0817)
Age = 10-19	-1.088***	-1.100***	-1.094***
	(0.0591)	(0.0591)	(0.0591)
Age = 30-39	1.040***	1.038***	0.9997***
	(0.0285)	(0.0285)	(0.0285)
Age = 40-49	1.973***	1.970***	1.951***
	(0.0260)	(0.0260)	(0.0260)
Age = 50-59	2.935***	2.925***	2.986***
	(0.0248)	(0.0248)	(0.0249)
Age = 60-69	4.029***	4.014***	4.154***
	(0.0244)	(0.0244)	(0.0245)
Age = 70-79	5.049***	5.040***	5.259***
	(0.0243)	(0.0243)	(0.0244)
Age = 80+	6.172***	6.231***	6.523***
	(0.0243)	(0.0243)	(0.0244)
Sex = Male		0.5246***	0.5296***
		(0.0041)	(0.0042)
Sex = Other		0.3811	0.6075
		(0.5610)	(0.5615)
Race = Asian			1.168***
			(0.0120)
Race = Black			0.6839***
			(0.0063)
Race = Hispanic			0.9556***
			(0.0058)
Race = Multiple/Other			0.2446***
			(0.0140)
Race = Native/PI			0.9924***
			(0.0180)
Fit statistics			
Observations	5,323,076	5,323,076	5,323,076
Squared Correlation	0.24461	0.25011	0.26402
Pseudo R^2	0.34602	0.35242	0.36748
BIC	1,675,710.2	1,659,349.3	1,620,833.6

IID standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: This table displays point estimates and standard errors from logistic regression models estimated using the CDC COVID-19 Case Surveillance Public Use Data (without geography), which has information on infected patients' demographics and health outcomes. The outcome variable is an indicator for whether a patient died as a result of COVID-19 infection. The excluded categories (baseline characteristics) are age 20-29, sex female, and race white. Observations with missing values for the outcome variable or any of the covariates are excluded.

Figure 3: Predicted Fatality Rate by Age and Sex



Notes: This plot shows the predicted probability of death conditional on infection for different demographic groups using parameter estimates from the logistic regression model in Table 1, Column 3. The probabilities are calculated for white individuals with each age group and sex. Predicted probabilities for ages 0-9 and 10-19, not displayed, are less than 0.001.

Figure 3 plots the predicted fatality rate by age and sex for white individuals, using the parameter estimates of the logistic regression model in Table 1, Column 3. The figure shows that the predicted fatality rates increase exponentially with age. It also illustrates the difference in predicted fatality rates across sexes, which becomes large (economically significant) at higher ages: for example, the gap between the predicted probability of death conditional on infection for men and that for women reaches almost eight percentage points for ages 70 to 79. Coupled with figure 2, this figure illustrates the sources of variation in the mortality risk measure we construct.

Appendix Table C.1 shows the parameter estimates of the logistic regression model including geographic characteristics. Although we do not use these models for predictions in the next part of the paper because of the coarse age bins available in the CDC data that includes geographic information, it is still useful to check the robustness of age-specific estimates to the inclusion of other potential determinants of fatality rate.

Although the age bins are different from those of our preferred specification, we see that the magnitudes of the coefficients on the age and sex variables remain similar across columns, despite including different sets of geographic characteristics. All coefficients are statistically significant at the 0.01 level. Although many of the geographic characteristics are correlated, there are some which have signs that are suggestive. The coefficient on physicians per 100,000 residents is negative, suggesting that the availability of medical care is correlated with lower fatality rates. The coefficients on percent below the poverty line, percent without a vehicle, and percent limited English speakers are negative, suggesting that higher financial, transportation, and language barriers to medical care are correlated with higher fatality rates. The coefficient on particulate matter (PM2.5) is also positive, suggesting that areas with higher pollution tend to have higher fatality rates. Suprisingly, the coefficients on percent smoking, obesity rate, and percent uninsured are negative, which suggests that areas with higher values of these variables tend to have lower fatality rates; because so many of these geographic factors are correlated, and our main goal of this portion is prediction rather than causal inference, we note but do not take much stock in these counterintuitive results. The takeaway from this robustness check is that the parameter estimates for the age and sex variables are not affected much when accounting for geographic determinants of fatality rate.

5 Empirics: Movement

Now, we use these estimates of mortality risk to study heterogeneous responses to the probability of death, as measured by how many venues like stores and restaurants people visited throughout the COVID-19 pandemic. We use variation in the incidence of the virus across time and space to identify the responsiveness of people of different age groups. We discuss how (under what assumptions) we can use this as revealed preference evidence for the distribution of the value of reduced mortality risk (VRMR) across ages.

5.1 Data

In this section, we describe our data on movement patterns and their determinants.

SafeGraph. Our empirical analysis of movement patterns primarily uses the Weekly Patterns dataset provided by SafeGraph, Inc.[9] SafeGraph provides aggregated information about visitors to approximately 6.8 million venues[10] across the United States from January 1, 2018 to the present. The data derive from a panel of approximately 45 million anonymized cell phones and track how often people visit venues and where they came from.[11]

In the data, each visitor–as identified by a unique cell phone–is assigned to a home Census block group based on the area in which the phone is observed most often during the hours between 6:00PM and 7:00AM. However, the Weekly Patterns dataset does not disclose individual-level visits. Instead, for a given week, it measures the number of visitors to a specific venue whose home is a particular Census block group.

We merge the Weekly Patterns dataset with the Core Places dataset to get additional information about each venue. Specifically, we get its geographic coordinates and its industry,

[9] SafeGraph provides this data free for academics at https://www.safegraph.com/academics.

[10] A venue is a specific site; for example, the gas station on LaSalle Street or the coffee shop on Hillsborough Road.

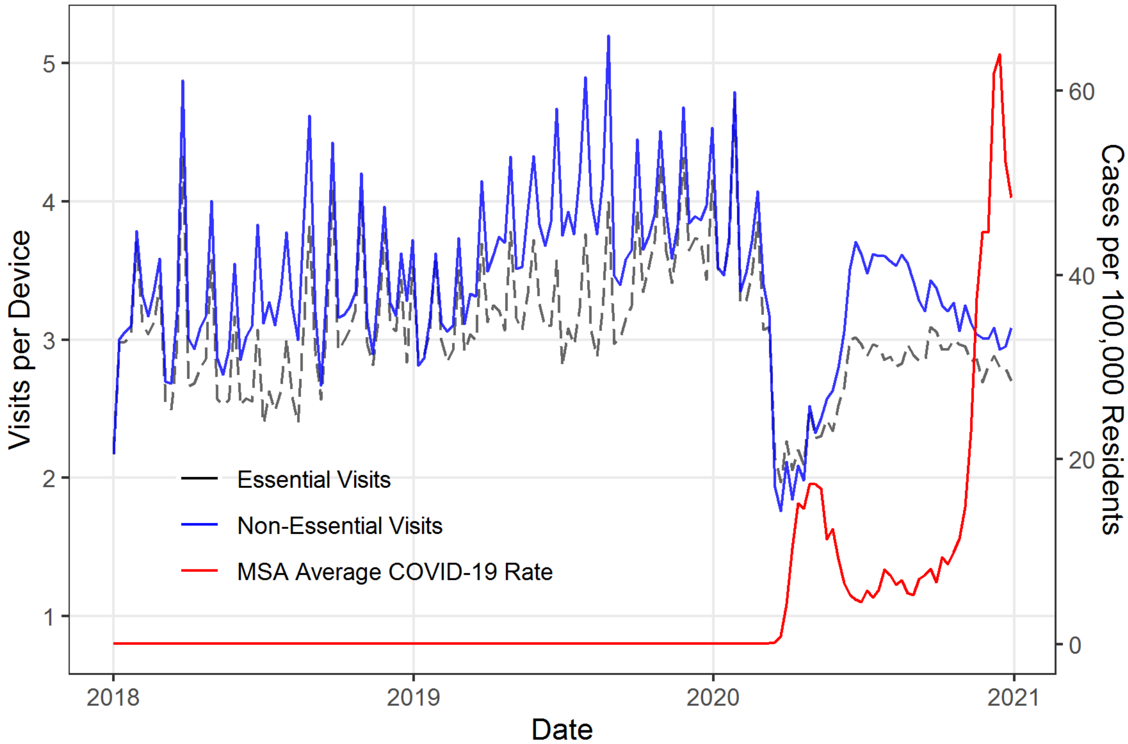
[11] These data are aggregated and anonymized by SafeGraph from cell-site location information (CSLI) microdata. CSLI data use records of when cell signals bounce off ("ping" at) cell towers, base stations, and antennae to approximate the location of an individual.

as determined by its North American Industry Classification System (NAICS) code. We classify venues into fifteen categories, which we then identify as essential or non-essential visits. The venue categories labeled essential are agriculture, industrial, transportation and wholesale, food stores, pharmacies and gas stations, offices, education, medical services, residential and elderly care, and other essential services. These include venues that serve basic needs, like food, healthcare, and education, and those where most visits would likely be for work rather than consumption. The venue categories labeled non-essential are retail, recreation and entertainment, hotels and lodging, restaurants, and other non-essential services. The specific types of venues that go into each category, as well as the number of locations of each sub-type tracked in the dataset across the United States, are listed in Appendix Table B.1. Note that we cannot distinguish between workers and visitors at non-essential venues like retail and restaurants. However, the data track the number of visitors to the venue in a given week, so these workers would only get counted once even if they visit multiple days of the week; this is a limitation of our data that we cannot overcome.

In this paper, we focus on the movements of people residing in the Philadelphia Metropolitan Statistical Area (MSA).[12] This restriction is made largely due to computational constraints; since 6.8 million venues are observed each week for more than four years, the full dataset is large (about 1.5 TB), and it would be difficult to estimate a regression model on the entire country. We also restrict our attention to visits from January 1, 2018 to December 31, 2020. We include a significant pre-period to capture the baseline movement levels when people face no COVID-19 infection-related mortality risk. We stop our analysis at the end of 2020 because of the introduction of vaccines. Vaccination status would introduce a highly endogenous determinant of movement levels. After this period, the decision to mitigate mortality risk becomes a higher dimensional problem, where the choice variables include both whether to move and whether to get the vaccine, which are not chosen independently.

[12] The Philadelphia-Camden-Wilmington MSA includes the following counties: New Castle County, Delaware; Cecil County, Maryland; Burlington, Camden, Gloucester, and Salem Counties, New Jersey; and Bucks, Chester, Delaware, Montgomery, and Philadelphia Counties, Pennsylvania.

Figure 4: Time Series of Essential and Non-Essential Visits in Philadelphia



Notes: This figure plots the time series of essential and non-essential visits in the Philadelphia Metropolitan Statistical Area (MSA) from January 1, 2018 to December 31, 2020. The number of visits in a given week is calculated by summing the number of visitors who reside in any of the counties in the MSA across all venues tracked in the SafeGraph data. The visits are normalized by the number of cell phones tracked in the MSA in a given week. Venues are categorized as essential and non-essential as described in Appendix B. The red line plots the MSA average COVID-19 infected rate. This is calculated by taking the mean of the eleven MSA counties' seven-day rolling average of new COVID-19 cases measured by the *New York Times* data.

We normalize the number of visits to venues of each category using the home panel summary, which counts the number of devices in the panel that reside in each block group. Our outcome variable for empirical analysis is the number of non-essential visits per cell phone tracked in a given block group and a given week. Figure 4 plots the weekly number of essential and non-essential visits per device in the Philadelphia metropolitan area. We see a significant dip in visits at the onset of the pandemic in March and April 2020 and another dip in late 2020 when the incidence of COVID-19 is high again.

American Community Survey. We supplement our information about movement patterns with demographic data from the 2015 to 2019 American Community Survey (ACS) 5-year estimates.[13] We observe block-group level demographic characteristics through counts of the number of people in various age bins.[14] We combine this information about age with that about the home block group of visitors to each venue in the SafeGraph data to infer age-specific movement patterns.

We also use a number of other block group characteristics measured by the ACS that may also determine movement levels. We use data on median household income, the share of households without internet access, and the share of essential workers,[15] which previous studies have established as salient determinants of movement. We also use measures of the percent of households below the poverty line, the percent of the population without health insurance coverage, the percent of the population with at least a bachelor's degree, and the percent of households that are renters.

To help measure the quality of an individual's outside option of staying home, we also use information about average household size, crowding (the percent of households with greater than one person per room), the percent of households without access to a vehicle, the percent of housing units lacking complete plumbing facilities, the percent of housing units lacking complete kitchen facilities, the percent of households living in mobile homes, and the percent of households living in group quarters or institutional living.

5.2 Model

In this section, we describe how we learn about elements of the movement-mortality risk tradeoff from our data. To complement our theoretical model, we develop an empirical model

[13] We obtain block group-level ACS demographic data from SafeGraph's Open Census Data, documented at https://docs.safegraph.com/docs/open-census-data.

[14] We also use tract-level data to approximate block group-level race distributions because the cross-tabulations of age, sex, and race are not available at the block group-level, as described in Appendix A.

[15] We define essential workers as individuals working in healthcare practitioners and technical occupations; service occupations, including healthcare support, protective service, food preparation and serving, building and grounds cleaning and maintenance occupations; natural resources, construction, and maintenance occupations; and production, transportation, and material moving occupations.

of movement decisions. We use individual-level potential outcomes to derive the appropriate specification for analyzing movement patterns aggregated to the block group level.

Realized Outcomes. Let Y be a random variable denoting the number of venues visited (by an individual i at time t). Let A be a discrete random variable for age group. Let M be a random variable for the mortality risk faced per unit of Y. As derived in the theoretical framework in equation 9 and estimated according to equation 19, this variable is calculated using the share of infected residents and predictions from the logit model of the probability of death conditional on exposure. Let X be a random vector of other observed determinants of movement (control variables), which could include time-variant and time-invariant individual characteristics and characteristics of the geography in which they live. Finally, let U represent unobserved determinants of movement. Equation 20 highlights that the realized movement level Y of a given individual is a function of the realized values of these other random variables and vectors.

$$Y = f(A,\ M,\ X,\ U) \tag{20}$$

Potential Outcomes. To discuss causality and treatment effects, we consider the potential outcomes $Y(a, m, x)$, which denote what an individual's movement level would have been under different values of the covariates than those that were realized: for example, if they were a different age or faced a different mortality risk. We parameterize these potential outcomes using a functional form assumption:

$$Y(a,\ m,\ x) = \sum_a 1_{[A=a]} \cdot \beta_a + \sum_a 1_{[A=a]} \cdot \delta_a \cdot m + \theta\ x + U \tag{21}$$

Assuming linearity of potential outcomes, as in equation 21, has a few important implications. First, although *levels* of movement can vary across individuals based on their values of x, this functional form implicitly assumes the individual-level treatment effects of mortality risk $\Delta(m_0 \rightarrow m_1)$ are homogeneous within age groups, as shown in equation 22.

$$\Delta(m_0 \rightarrow m_1) = Y(a', m_1, x') - Y(a', m_0, x') = (\delta_{a'})(m_1 - m_0) \tag{22}$$

It also allows δ to be a sufficient statistic to capture the heterogeneity in movement-mortality risk trade-offs across different age groups. Recall that the focus of our analysis is how dy/dm varies by age. Derived using the functional form given by equation 21, equation 23 shows that δ is sufficient to determine this relationship: it is the gradient of movement with respect to mortality risk.

$$\left.\frac{dy}{dm}\right|_{A=a} = \delta_a \tag{23}$$

Aggregating to the Block Group. Including subscripts, the realized outcome Y_{ijt} for individual i residing in block group j moving during week t will take the form of equation 24.

$$Y_{ijt} = \sum_a 1_{[Age_i=a]} \cdot \beta_a + \sum_a 1_{[Age_i=a]} \cdot \delta_a \cdot M^a_{jt} + \theta X_{ijt} + U_{ijt} \tag{24}$$

To aggregate to block group level movement patterns, we sum over individuals and divide by the block group population to yield equation 25.[16] $\bar{Y}$ measures the number of visits per capita (or, per cell phone in our panel) of individuals residing in a block group. s^a represents the share of the block group with age a.

$$\bar{Y}_{jt} = \sum_a \beta_a \cdot s^a_j + \sum_a \delta_a \cdot s^a_j \cdot M^a_{jt} + \theta \bar{X}_{jt} + \bar{U}_{jt} \tag{25}$$

Since our fatality rate estimates depend on not just age but also sex and race, the process to aggregate to the block group is slightly more complex and described in Appendix A.

[16]As an intermediate step, note that

$$\sum_i Y_{ijt} = \sum_a \beta_a \cdot n^a_j + \sum_a \delta_a \cdot n^a_j \cdot M^a_{jt} + \theta \sum_i X_{ijt} + \sum_i U_{ijt}$$

Regression Specification. Now, combining equations 19 and 25, we derive our regression specification 26.

$$Y_{jt} = \sum_a \beta_a \cdot s_j^a + \sum_a \delta_a \cdot s_j^a \cdot \hat{M}_{jt}^a + \theta X_j + \psi_j + \xi_t + \epsilon_{jt} \tag{26}$$

Here, Y_{jt} is the amount of non-essential movement, measured in visits per capita. s_j^a is the share of block group j with age a. $\hat{M}_{jt}^a$ is the mortality risk faced by an individual of age a in block group j at time t, calculated according to equation A.6. X_j is a vector of time-invariant block group characteristics such as income and internet access, whose inclusion is motivated by the literature on heterogeneity in mobility responses to the pandemic. η_j are county fixed effects and ξ_t are week fixed effects, which control for time-varying determinants of visits like movement restriction policies and weather. ϵ_{jt} is an idiosyncratic error term.

Value of Reduced Mortality Risk. In the regression specified by equation 26, $\hat{\delta}_a$ gives an estimate of the gradient of movement with respect to mortality risk for each age group a. We are ultimately interested in how the value of reduced mortality risk (VRMR) varies by age, since this is a key parameter in cost-benefit analysis. VRMR measures the marginal willingness to pay for a marginal reduction in the probability of death. Recall from equation 17 that VRMR is the product of the gradient of movement with respect to mortality risk (equivalently, the marginal rate of substitution between movement and the probability of death)[17] and the marginal willingness to pay for movement. This relationship is reflected in equation 27.

$$V\widehat{RM}R(a) \propto \hat{\delta}_a \cdot MWTP_y(a) \tag{27}$$

After estimating $\hat{\delta}_a$, we consider different assumptions about the shape of $MWTP_y(a)$ and see what we can conclude about $VRMR(a)$.

[17]Note that we estimate dy/dm and $M \propto p^D$, so $\hat{\delta}$ is proportional to dy/dp^D rather than being equal to.

5.3 Results

Regression Estimates. Table 2 shows ordinary least squares (OLS) estimates of the parameters relating age and mortality risk to movement, $\hat{\beta}_a$ and $\hat{\delta}_a$. The response variable is the number of non-essential visits per cell phone for a block group j in a week t. We do not consider the share of residents aged zero through nine years as they are unlikely to have phones that are tracked in the data. Columns 1 and 2 do not include any fixed effects, while Column 3 includes week fixed effects. Column 2 includes the infection rate, measured as the county seven-day rolling average of new COVID-19 cases, as a control variable. Although this factor is included in the calculation of the mortality risk variable, it may also affect movements directly.

The coefficients on the age shares in Columns 1 and 2 are estimates of β_a and can be interpreted as the baseline level of non-essential visits made by people of that age group, since there is no excluded category and no constant in the regression. We see that people aged 10-19 have particularly high baseline movement levels, those aged 20-29 and 30-39 tend to move somewhat less, those aged 50-59 also have high baseline levels, and beyond this age, baseline movement tends to decrease. The coefficients on the age shares in Column 3 are not as straightforward to interpret because of the inclusion of week fixed effects. All age share coefficients are statistically significant at the 0.01 level.

The interaction terms between age shares and mortality risk are estimates of δ_a and can be interpreted as the gradient of movement with respect to mortality risk, as shown in equation 23. In columns 1 and 2, the coefficients are all negative and monotonically decreasing in magnitude with age; they are also all statistically significant at the 0.01 level. This suggests that people of all ages move less when they face higher mortality risk, and that younger people are more responsive to mortality risk than older people. In column 3, which includes week fixed effects, the coefficients on the interaction terms for ages 30-39 and 40-49 are insignificant. However, all other coefficients follow the same pattern: they are negative and decreasing in magnitude with age.

Table 2: Movement Parameter Estimates: Age and Mortality Risk

Dependent Variable:	Non-Essential Visits		
Model:	(1)	(2)	(3)
Variables			
Share 10-19 Years	4.494***	4.491***	1.726***
	(0.0208)	(0.0208)	(0.1588)
Share 20-29 Years	3.468***	3.466***	1.040***
	(0.0165)	(0.0165)	(0.1342)
Share 30-39 Years	3.953***	3.952***	0.7695***
	(0.0206)	(0.0206)	(0.0872)
Share 40-49 Years	4.211***	4.206***	1.442***
	(0.0277)	(0.0277)	(0.0955)
Share 50-59 Years	4.315***	4.300***	1.823***
	(0.0246)	(0.0246)	(0.0960)
Share 60-69 Years	3.844***	3.837***	1.430***
	(0.0268)	(0.0268)	(0.0939)
Share 70-79 Years	3.874***	3.880***	1.649***
	(0.0353)	(0.0353)	(0.1123)
Share 80+ Years	3.367***	3.390***	1.025***
	(0.0360)	(0.0360)	(0.1149)
Mortality Risk × Share 10-19 Years	-20.88***	-45.82***	-15.78***
	(2.314)	(2.524)	(5.074)
Mortality Risk × Share 20-29 Years	-9.904***	-16.06***	-6.806***
	(0.6326)	(0.6798)	(1.553)
Mortality Risk × Share 30-39 Years	-1.283***	-5.168***	0.8324
	(0.2970)	(0.3360)	(0.9444)
Mortality Risk × Share 40-49 Years	-0.8622***	-2.290***	-0.2381
	(0.1546)	(0.1650)	(0.2970)
Mortality Risk × Share 50-59 Years	-0.3839***	-0.9061***	-0.2451**
	(0.0539)	(0.0579)	(0.0963)
Mortality Risk × Share 60-69 Years	-0.2442***	-0.4098***	-0.1823***
	(0.0202)	(0.0213)	(0.0251)
Mortality Risk × Share 70-79 Years	-0.1360***	-0.2203***	-0.1106***
	(0.0116)	(0.0121)	(0.0150)
Mortality Risk × Share 80+ Years	-0.0992***	-0.1507***	-0.0776***
	(0.0061)	(0.0064)	(0.0079)
Infection Rate		0.0152***	
		(0.0006)	
Fixed-effects			
Week			Yes
Fit statistics			
Observations	665,627	665,627	665,627
R^2	0.01524	0.01614	0.25968
Within R^2			0.01402

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: This table displays ordinary least squares (OLS) estimates of the parameters relating age and mortality risk to non-essential visits. The dependent variable is the number of non-essential visits per cell phone tracked in a block group in a given week. Non-essential visits are measured using SafeGraph data and classified based on NAICS code as described in Appendix B. Age shares are derived from American Community Survey 2015-2019 5-year estimates. The mortality risk measure is constructed as described in Appendix A. In columns one and two, the coefficients on the age shares estimate the baseline level of non-essential visits for that age in a given week; there is no excluded group or constant. In column three, there are week fixed effects, so the same interpretation does not hold. In all columns, the coefficients on the age shares interacted with mortality risk estimate the responsiveness a person of that age to mortality risk.

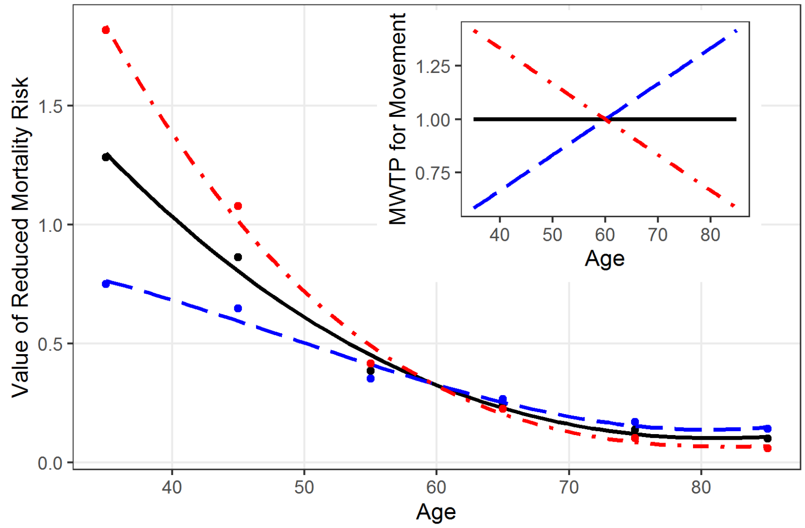
Appendix Table C.2 checks the robustness of these findings by running alternative specifications. Column 1 includes both week and county fixed effects. The interaction terms for ages 10-19 through 50-59 become statistically insignificant, so we cannot identify the age pattern in this specification. Note that the mortality risk measure uses county-level COVID-19 infection rates in its construction, and so including county fixed effects might be restricting some of the variation that is identifying the coefficients in the other specifications. However, in lieu of fixed effects but in an attempt to still account for potential geographic factors, we include a vector of block group characteristics derived from the ACS estimates in columns 3 through 5. Columns 2 and 5 also include the number of essential visits as a control. This reflects that people who already move for an essential visit may engage in trip chaining by also engaging in non-essential movement while out of the house. In Columns 2 and 5, the interaction terms for ages 40-49 and 50-59 are not statistically significant. However, all coefficients that are significant follow the pattern that responsiveness to mortality risk is decreasing in age. Columns 3 and 5 have coefficients that monotonically decrease in magnitude, which again suggests that older people tend to be less responsive to mortality risk. In summary, although in some specifications some coefficients are not statistically significant, the pattern seems to hold even under the inclusion of a variety of other controls.

Comparative Statics. We now use these estimates of the derivative of movement with respect to mortality risk to discuss what our findings imply about the shape of the value of reduced mortality risk (VRMR) across age groups, using the relationship presented in equation 27. We have estimates of the the gradient of movement with respect to mortality risk. Now, in a comparative statics exercise, we see what we can conclude about VRMR when making different assumptions about the shape of the way the marginal willingness to pay for movement varies by age. We claim that under weak assumptions, our results suggest that VRMR is either monotonically decreasing or inverse U-shaped (eventually decreasing) in age.

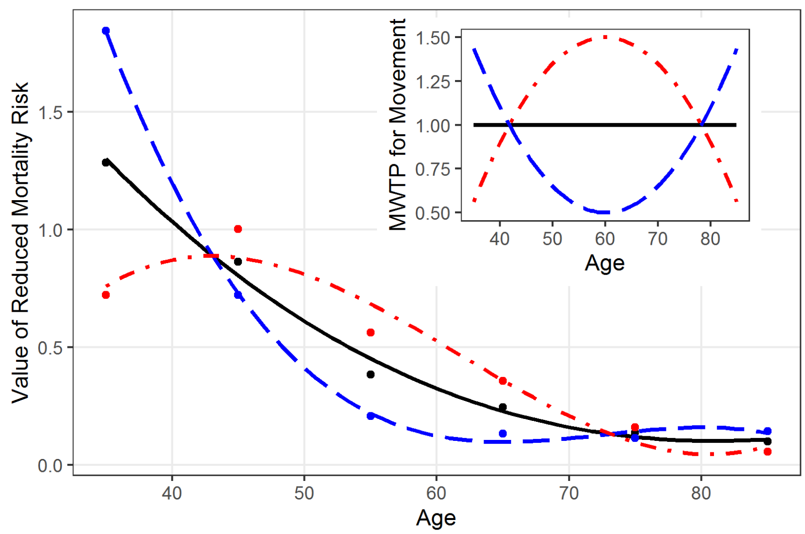
Figure 5 plots five different potential shapes for the MWTP for movement–constant,

Figure 5: Comparative Statics: Value of Reduced Mortality Risk

(a) Increasing and Decreasing MWTP



(b) U-Shaped and Inverse U-Shaped MWTP



Notes: This figure plots the relative magnitude of the value of reduced mortality risk across age groups under different shapes of the marginal willigness to pay (MWTP) for movement. The inset graphs show the assumed shapes of the MWTP curves. In Panel A, we allow the curve to be linearly increasing (in blue) and linearly decreasing (in red). In Panel B, we allow the curve to be U-shaped (in blue) and inverse U-shaped (in red). Both panels also include the case when MWTP is constant as a reference. The larger graphs show the value of reduced mortality risk curves under the assumed shapes of the MWTP curves. The point estimates for each age are calculated by multiplying the coefficient estimates from Table 2, Column 1 by the value of the assumed MWTP curve at that age. These points are connected using smooth third-order polynomial approximations.

increasing, decreasing, U-shaped, and inverse U-shaped–in the inset graphs and shows the shape of the VRMR curve under these assumptions in the larger graphs. The point estimates of the value of reduced mortality risk for each age under the assumed MWTP shape are calculated by multiplying the coefficient estimates from Table 2, Column 1, by the value of the assumed MWTP curve at that age. To visualize the curves' shapes, these points are connected using smooth third-order polynomial approximations.

Panel 5a shows constant (in black), linearly increasing (in blue), and linearly decreasing (in red) forms of the MWTP for movement in the inset graph and plots the resulting VRMR curves in the larger graph. If MWTP is constant across ages, VRMR maintains the shape of the marginal rate of substitution between mortality risk and movement, estimated by $\hat{\delta}_a$; in this case, it is monotonically decreasing. If MWTP is linearly increasing in age, the slope of the VRMR curve is flatter than the slope in the case of constant MWTP. Eventually, if MWTP is sufficiently increasing in age, the VRMR curve would become flat and, if extreme enough, eventually increasing. However, the MWTP for movement for 70-79 year olds would need to be more than nine times the MWTP for 30-39 year olds for them to have the same VRMR, which is unrealistic, so we conclude that under this case VRMR would still be monotonically decreasing. If MWTP is linearly decreasing in age, the slope of the VRMR curve is steeper than the slope in the case of constant MWTP. In this case, no matter the slope, the VRMR will stay decreasing in age.

Panel 5b shows U-shaped (in blue) and inverse U-shaped (in red) MWTP curves and the resulting VRMR curves. If MWTP is U-shaped, the VRMR curve remains decreasing in age, but becomes more convex than under constant MWTP. If MWTP is inverse U-shaped, the VRMR curve begins to take an inverse-U shape, especially as the curvature of the MWTP curve increases. However, the VRMR is still eventually decreasing. The literature review highlights that previous studies have found VRMR to be inverse U-shaped; our results are compatible with this finding. After examining these five cases, we conclude that under weak assumptions, our results imply that VRMR is monotonically or eventually decreasing in age.

6 Discussion

Policy Implications. Previous studies have shown that how the willingness to pay for risk reduction varies with age is theoretically ambiguous. Over the life cycle, life expectancy shortens but economic resources also vary. Our empirical results suggest that older people are less responsive to mortality risk and have a lower willingness to pay for marginal reductions in mortality risk. Specifically, under weak assumptions, our findings indicate that the value of reduced mortality risk (VRMR) is monotonically or eventually decreasing in age.

This finding is consistent with previous studies in the VSL literature which have found that the value of reduced mortality risk takes an inverse U-shape: the willingness to pay for reductions in the probabiliy of death initially increases and eventually decreases in age. Most previous studies focusing on how VRMR varies across age groups have looked to evidence from labor market tradeoffs between wages and on-the-job fatality risk, examining how much extra pay is required to induce people to accept jobs in which workers face higher risk. One limitation of using the labor market setting to analyze age patterns in VRMR is that most workers retire around age 65. Thus, most of the identifying variation is from the decisions of relatively young people, so estimates may not generalize or be precise for the elderly. Although there may be some similar concerns in our study due to different levels of mobile phone adoption across age groups, we argue that conducting validation studies in alternative settings is needed to support and extend the VSL literature. The setting and approach of this paper is novel and contributes to the body of evidence that VRMR decreases with age.

The finding that VRMR is monotonically or eventually decreasing in age has policy implications for cost-benefit analysis. This estimate is an important parameter that frequently accounts for the majority of quantified benefits in policies and programs like safety regulations and pollution remediation designed to mitigate adverse health outcomes. There remains a political and ethical controversy among policymakers about whether to value reductions in mortality risk to older people differently from those to younger age groups, as evidenced in controversial debates about the use of a "senior discount" in cost-benefit analysis. However, research studying the willingness to pay for mortality risk reductions can instruct policymakers who seek to prioritize risk reduction efforts towards specific populations who value it most. Showing and estimating heterogeneity in VRMR allows for more precise cost-benefit analysis and more informed policymaking.

Limitations and Extensions. There are several limitations to this analysis and extensions that could be conducted. The most obvious of these is extending the geographic scope of the empirical analysis of movement patterns beyond the Philadelphia metropolitan area. This is feasible, albeit computationally challenging. Just because these results hold in Philadelphia do not mean they will hold across the entire United States. Including other areas, such as the fifty largest Metropolitan Statistical Areas (MSAs), would show the robustness of these findings and allow us to characterize any heterogeneity across cities and regions.

Another feasible improvement would be to use microdata tracking individual cell phones instead of examining aggregate movement patterns. While not free, these data are available and have been used in other studies like Couture et al. (2021). However, despite being able to observe individual-level movement decisions, we would still need to make assumptions about the age of the device owner because no data broker includes personal demographic information due to confidentiality concerns.

Further work could study the role of heterogeneous risk misperceptions and household externalities in movement decisions. In this paper we have assumed that people correctly perceive the mortality risk they face, at least on average. This is likely not the case, so relaxing this assumption could be a fruitful avenue for future research. Also, individuals' movement decisions may be based on not only the mortality risk they would face but also the risk other members of their household would incur due to their movement. For example, a young person living with an elderly family member may be careful to limit potential exposure not for their own sake but because others in their household face a high probability of death conditional on infection.

The last limitation worth mentioning is that we do not characterize how the choice set of venues to visit may change over time. This may be a concern because of movement restrictions policies like stay-at-home orders and store closures. Our concern would be that an individuals' choice to move less would reflect this constraint rather than their preferences. We include time fixed effects in our empirical analysis in order to account for such changes, but further work could try to more rigorously separate the role of preferences from that of constraints on choice sets.

7 Conclusion

This paper set out to study whether responsiveness to mortality risk varies by age. We use data from health outcomes and movement patterns during the COVID-19 pandemic to show evidence that it does. Specifically, we find that older people have a lower willingness to pay for marginal reductions in the probability of death.

We develop a theoretical model characterizing movement decisions and find that people optimize by equating the marginal utility of moving with the product of the marginal probability of infection and the expected disutility of infection (and potentially death). We also highlight potential sources of heterogeneity in movement decisions.

Recognizing that probability of death is jointly determined by the probability of infection and the probability of death conditional on infection, we use data on the geographic incidence of COVID-19 and the health outcomes of infected patients to construct an empirical measure of mortality risk. In the process, we examine demographic and geographic heterogeneity in fatality rates, finding that old, non-white, or male patients tend to face adverse health outcomes from COVID-19 infection more frequently than young, white, and female patients.

Using this measure, we study the responsiveness of people of different age groups to infection-related mortality risk. We use data from a panel of cell phones tracking visits to venues such as stores and restaurants before and during the COVID-19 pandemic to estimate a regression model that quantifies the gradient of movement with respect to mortality risk. We find that older people's movements tend to be less elastic to changes in mortality risk. Then we use comparative statics to show that this implies that the willingness to pay for marginal risk reductions is eventually decreasing in age.

Our study contributes novel evidence and methods to the literature examing the value of statistical life (VSL) and sheds light on heterogeneity in a key parameter for the cost-benefit analysis of many government policies and regulations. Policymakers can weigh this evidence with other political and ethical concerns when evaluating how much to value risk reductions.

References

Aldy, J. E. and Viscusi, W. K. (2007). Age differences in the value of statistical life: Revealed preference evidence. *Review of Environmental Economics and Policy*, 1(2):241–260.

Aldy, J. E. and Viscusi, W. K. (2008). Adjusting the value of a statistical life for age and cohort effects. *The Review of Economics and Statistics*, 90(3):573–581.

Alexander, D. and Karger, E. (2020). Do stay-at-home orders cause people to stay at home? Effects of stay-at-home orders on consumer behavior. *Federal Reserve Bank of Chicago Working Paper Series*, WP 2020-12.

Alfaro, L., Faia, E., Lamersdorf, N., and Saidi, F. (2020). Social interactions in pandemics: Fear, altruism, and reciprocity. *National Bureau of Economic Research Working Paper Series*, No. 27134.

Allcott, H., Boxell, L., Conway, J., Gentzkow, M., Thaler, M., and Yang, D. (2020a). Polarization and public health: Partisan differences in social distancing during the coronavirus pandemic. *Journal of Public Economics*, 191:104254.

Allcott, H., Boxell, L., Conway, J. C., Ferguson, B. A., Gentzkow, M., and Goldman, B. (2020b). What explains temporal and geographic variation in the early US coronavirus pandemic? *National Bureau of Economic Research Working Paper Series*, No. 27965.

Ashenfelter, O. (2006). Measuring the value of a statistical life: Problems and prospects. *The Economic Journal*, 116(510):C10–C23.

Banzhaf, H. S. (2014). Retrospectives: The Cold-War origins of the Value of Statistical Life. *The Journal of Economic Perspectives*, 28(4):213–226.

Barrios, J. M. and Hochberg, Y. (2020). Risk perception through the lens of politics in the time of the COVID-19 pandemic. *National Bureau of Economic Research Working Paper Series*, No. 27008.

Blomquist, G. C. (2004). Self-protection and averting behavior, values of statistical lives, and benefit cost analysis of environmental policy. *Review of Economics of the Household*, 2(1):89–110.

Broome, J. (1978). Trying to value a life. *Journal of Public Economics*, 9(1):91–100.

Cameron, T. A. (2010). Euthanizing the value of a statistical life. *Review of Environmental Economics and Policy*, 4(2):161–178.

Canning, D., Karra, M., Dayalu, R., Guo, M., and Bloom, D. E. (2020). The association between age, COVID-19 symptoms, and social distancing behavior in the United States. *medRxiv*, page 2020.04.19.20065219.

Chang, S., Pierson, E., Koh, P. W., Gerardin, J., Redbird, B., Grusky, D., and Leskovec, J. (2021). Mobility network models of COVID-19 explain inequities and inform reopening. *Nature*, 589(7840):82–87.

Chiou, L. and Tucker, C. (2020). Social distancing, internet access and inequality. *National Bureau of Economic Research Working Paper Series*, No. 26982.

Couture, V., Dingel, J. I., Green, A., Handbury, J., and Williams, K. R. (2021). Measuring movement and social contact with smartphone data: a real-time application to COVID-19. *Journal of Urban Economics*, page 103328.

Ehrlich, I. and Yin, Y. (2005). Explaining diversities in age-specific life expectancies and values of life saving: A numerical analysis. *Journal of Risk and Uncertainty*, 31(2):129–162.

Engle, S., Stromme, J., and Zhou, A. (2020). Staying at home: Mobility effects of COVID-19. *Working Paper*.

Gao, S., Rao, J., Kang, Y., Liang, Y., Kruse, J., Dopfer, D., Sethi, A. K., Mandujano Reyes, J. F., Yandell, B. S., and Patz, J. A. (2020). Association of mobile phone location data indications of travel and stay-at-home mandates with COVID-19 infection rates in the US. *JAMA Network Open*, 3(9):e2020485–e2020485.

Glaeser, E. L., Jin, G. Z., Leyden, B. T., and Luca, M. (2021). Learning from deregulation: The asymmetric impact of lockdown and reopening on risky behavior during COVID-19. *Journal of Regional Science*, 61(4):696–709.

Goolsbee, A. and Syverson, C. (2020). Fear, lockdown, and diversion: Comparing drivers of pandemic economic decline 2020. *National Bureau of Economic Research Working Paper Series*, No. 27432.

Gosak, M., Kraemer, M. U. G., Nax, H. H., Perc, M., and Pradelski, B. S. R. (2021). Endogenous social distancing and its underappreciated impact on the epidemic curve. *Scientific Reports*, 11(1):3093.

Hammitt, J. K. (2000). Valuing mortality risk: Theory and practice. *Environmental Science & Technology*, 34(8):1396–1400.

Hammitt, J. K. (2020). Valuing mortality risk in the time of COVID-19. *Journal of Risk and Uncertainty*, 61(2):129–154.

Heinzerling, L. (2000). The rights of statistical people. *Harvard Environmental Law Review*, 24(1):189.

Jay, J., Bor, J., Nsoesie, E. O., Lipson, S. K., Jones, D. K., Galea, S., and Raifman, J. (2020). Neighbourhood income and physical distancing during the COVID-19 pandemic in the United States. *Nature Human Behaviour*, 4(12):1294–1302.

Johansson, P.-O. (1996). On the value of changes in life expectancy. *Journal of Health Economics*, 15(1):105–113.

Johansson, P.-O. (2002). On the definition and age-dependency of the value of a statistical life. *Journal of Risk and Uncertainty*, 25(3):251–263.

Kavanagh, N. M., Goel, R. R., and Venkataramani, A. S. (2021). County-level socioeconomic and political predictors of distancing for COVID-19. *American Journal of Preventive Medicine*, 61(1):13–19.

Kniesner, T. J., Viscusi, W. K., and Ziliak, J. P. (2006). Life-cycle consumption and the age-adjusted value of life. *Contributions in Economic Analysis & Policy*, 5(1).

Lamb, M. R., Kandula, S., and Shaman, J. (2021). Differential COVID-19 case positivity in New York City neighborhoods: Socioeconomic factors and mobility. *Influenza and Other Respiratory Viruses*, 15(2):209–217.

Mongey, S., Pilossoph, L., and Weinberg, A. (2020). Which workers bear the burden of social distancing? *National Bureau of Economic Research Working Paper Series*, No. 27085.

Reluga, T. C. (2010). Game theory of social distancing in response to an epidemic. *PLOS Computational Biology*, 6(5):e1000793.

Rizaldi, A. A., Xie, S., Hubbard, R. A., and Himes, B. E. (2022). Neighborhood characteristics and COVID-19 incidence and mortality in Southeastern Pennsylvania. *AMIA Joint Summits on Translational Science Proceedings*, pages 429–438.

Schelling, T. (1968). The life you save may be your own. In Chase, S. B., editor, *Problems in Public Expenditure Analysis*, pages 127–62. Brookings Institution, Washington, DC.

Simon, N. B., Dockins, C., Maguire, K. B., Newbold, S. C., Krupnick, A. J., and Taylor, L. O. (2019). What's in a name? A search for alternatives to "VSL". *Review of Environmental Economics and Policy*, 13(1):155–161.

Smith, V. K., Evans, M. F., Kim, H., and Taylor, D. H. (2004). Do the near-elderly value mortality risks differently? *The Review of Economics and Statistics*, 86(1):423–429.

Sun, Z., Di, L., Sprigg, W., Tong, D., and Casal, M. (2020). Community venue exposure risk estimator for the COVID-19 pandemic. *Health & Place*, 66:102450.

Vandenbroucke, G. (2020). Endogenous social distancing in an epidemic. *Federal Reserve Bank of St. Louis Working Paper*, (2020-013).

Viscusi, W. K. (2009a). The devaluation of life. *Regulation & Governance*, 3(2):103–127.

Viscusi, W. K. (2009b). Reply to the comments on "The devaluation of life". *Regulation & Governance*, 3(3):306–309.

Viscusi, W. K. (2012). What's to know? Puzzles in the literature on the Value of Statistical Life. *Journal of Economic Surveys*, 26(5):763–768.

Viscusi, W. K. and Aldy, J. E. (2003). The value of a statistical life: A critical review of market estimates throughout the world. *Journal of Risk and Uncertainty*, 27(1):5–76.

Weill, J. A., Stigler, M., Deschenes, O., and Springborn, M. R. (2020). Social distancing responses to COVID-19 emergency declarations strongly differentiated by income. *Proceedings of the National Academy of Sciences*, 117(33):19658.

Yechezkel, M., Weiss, A., Rejwan, I., Shahmoon, E., Ben-Gal, S., and Yamin, D. (2021). Human mobility and poverty as key drivers of COVID-19 transmission and control. *BMC Public Health*, 21(1):596.

A Appendix: Mortality Risk Measure

In this appendix, we explain the calculations behind the measure of mortality risk. In our calculations we use three pieces of information to calculate the mortality risk faced by a person of a specific age in a given census block group: (1) the demographic composition of the block group, specifically the breakdowns of age, race, and sex; (2) the demographic-specific estimates of fatality rates from the logistic regression models of death conditional on infection; and (3) the county-level infection rates.

Motivation. Recall from equations 24 and 25 that aggregating individual-level potential outcomes to the block group level gives the form

$$\bar{Y}_{jt} = \sum_a \beta_a \cdot s^a_j + \sum_a \delta_a \cdot \underbrace{s^a_j \cdot M^a_{jt}}_{\text{calculated}} + \theta \bar{X}_{jt} + \bar{U}_{jt}$$

where $\bar{Y}_{jt}$ is the average movement level of block group j, s^a_j is the share of block group j residents with age a, M^a_{jt} is a measure of the mortality risk faced by an individual of age a in block group j at time t. Here, we explain how we calculate the value of $s^a \cdot M^a$.

Calculations. This value has two components. The first component represents infection rate. The probability of infection is proxied for using the share infected s^I, measured by the county-level seven-day rolling average of new COVID-19 cases. The second component represents fatality rate. We take the logit estimates $\hat{f}^{a,s,r}$ of the probability of death conditional on infection for an individual of age a, sex s, and race r and integrate over the empirical distribution of block group demographics, represented by cumulative distribution function $F_{a,r,s}$.

$$s_a \cdot M_a = \underbrace{s^I}_{\text{infection rate}} \cdot \underbrace{\int_s \int_r \hat{f}^{a,s,r} dF_{a,r,s}}_{\text{fatality rate weighted by demographics}} \tag{A.1}$$

We write out this integration using summation notation to show how we conduct the calculations. $s^{a,s,r}$ represents the share of the block group that has age a, sex s, and race r; effectively, it is a joint probability. Similarly, $s^{a,s} = \sum_r s^{a,s,r}$ represents the share of the block group with age a and sex s. $s^{r|a,s}$ is the share of the population of the block group of age a and sex s who identify as race r; effectively, it is a conditional probability. Note that age share $s^a = \sum_s \sum_r s^{a,s,r}$, so the age share portion of $s^a \cdot M^a$ is incorporated within this integral expression.

$$
\begin{aligned}
&\int_s \int_r \hat{f}^{a,s,r} dF_{a,s,r} \\
&= \sum_s \sum_r s^{a,s,r} \hat{f}^{a,s,r} && \text{(A.2)} \\
&= \sum_s \sum_r s^{a,s} \cdot s^{r|a,s} \cdot \hat{f}^{a,s,r} && \text{(A.3)}
\end{aligned}
$$

ACS estimates report counts, not shares. We show how we calculate these shares from counts n.

$$
s^{r|a,s} = \frac{n^{a,s,r}}{n^{a,s}} = \frac{n^{a,s,r}}{\sum_{r'} n^{a,s,r'}} \text{ and } s^{a,s} = \frac{n^{a,s}}{n^{total}} = \frac{n^{a,s}}{\sum_{a',s'} n^{a',s'}} \quad \text{(A.4)}
$$

We observe $n_{bg}^{a,s}$ at the block group level, so we can calculate $s^{a,s}$. However, we only observe $n_t^{a,s,r}$ at the tract level; ACS estimates of this statistic are only published at this larger geography because of concerns about privacy and precision. Therefore we approximate $s_{bg}^{r|a,s}$ with $s_t^{r|a,s}$.

As a further complication, the age bins do not line up perfectly in the CDC and ACS data. Although the block group-level ACS data can match the age groups of the fatality rate estimates when calculating $s^{a,s}$, this is not the case for the tract-level ACS data. The age bins for the fatality rate estimates $\hat{f}^{a,s,r}$ are 0-9, 10-19, 20-29, 30-39, 40-49, 50-59, 60-69, 70-79, and 80+. The age bins for tract-level demographic counts $n_t^{a,s,r}$ are 0-5, 5-9, 10-14, 15-17, 18-19, 20-24, 25-29, 30-34, 35-44, 45-54, 55-64, 65-74, 75-84, 85+. Because of the imperfect age bin overlap, we use approximations:

$$s_{bg}^{r|a,s} \equiv \begin{pmatrix} s_{bg}^{r|a=0-9,s} \\ s_{bg}^{r|a=10-19,s} \\ s_{bg}^{r|a=20-29,s} \\ s_{bg}^{r|a=30-39,s} \\ s_{bg}^{r|a=40-49,s} \\ s_{bg}^{r|a=50-59,s} \\ s_{bg}^{r|a=60-69,s} \\ s_{bg}^{r|a=70-79,s} \\ s_{bg}^{r|a=80+,s} \end{pmatrix} \approx \begin{pmatrix} s_{t}^{r|a=0-9,s} \\ s_{t}^{r|a=10-19,s} \\ s_{t}^{r|a=20-29,s} \\ s_{t}^{r|a=25-44,s} \\ s_{t}^{r|a=35-54,s} \\ s_{t}^{r|a=45-64,s} \\ s_{t}^{r|a=55-74,s} \\ s_{t}^{r|a=65-84,s} \\ s_{t}^{r|a=75+,s} \end{pmatrix} \equiv \tilde{s}_t^{r|a,s} \tag{A.5}$$

In summary, the estimate of $s_a \cdot M_a$ is a calculated according to equation A.6 using the county infection rate s_{cty}^I, the block group-level joint distribution of age and sex $s_{bg}^{a,s}$, the tract-level conditional distribution of race given age and sex (using approximate age bins for those that do not line up) $\tilde{s}_t^{r|a,s}$, and the fatality rate estimates for age, sex, and race groups $\hat{f}^{a,s,r}$.

$$\sum_a \delta_a \cdot s_a \cdot M_a = \sum_a \delta_a \cdot s_{cty}^I \cdot \sum_s \sum_r s_{bg}^{a,s} \cdot \tilde{s}_t^{r|a,s} \cdot \hat{f}^{a,s,r} \tag{A.6}$$

B Appendix: NAICS Code Classifications

Table B.1 shows the categorization of venue types by NAICS code. It also displays our classificiations of essential and non-essential venue categories. The right-most column displays the number of venues of each sub-type that are tracked in the SafeGraph data across the United States.

Table B.1: NAICS Code Classifications.

Category	Classification	NAICS	Description	Number of Venues
Agriculture	Essential	11	Agriculture, Forestry, Fishing and Hunting	1,627
Industrial	Essential	21	Mining	34
		22	Utilities	9,069
		23	Construction	54,915
		31	Manufacturing	45,259
		32	Manufacturing	19,956
		33	Manufacturing	28,869
		562	Waste Management and Remediation Services	16,308
Transportation and wholesale	Essential	42	Wholesale Trade	76,906
		48	Transportation	48,874
		49	Warehousing	53,456
Food stores	Essential	4451	Grocery Stores	138,775
		4452	Specialty Food Stores	33,066
Retail	Non-essential	4411	Automobile Dealers	56,983
		4412	Other Motor Vehicle Dealers	26,032
		4413	Automotive Parts, Accessories, and Tire Stores	79,233
		4421	Furniture Stores	39,681
		4422	Home Furnishings Stores	30,501
		4431	Electronics and Appliance Stores	44,873
		4441	Building Material and Supplies Dealers	51,628
		4442	Lawn and Garden Equipment and Supplies Stores	25,226
		4453	Beer, Wine, and Liquor Stores	32,539
		4481	Clothing Stores	95,230
		4482	Shoe Stores	21,139
		4483	Jewelry, Luggage, and Leather Goods Stores	32,507
		4511	Sporting Goods, Hobby, and Musical Instrument Stores	77,357
		4512	Book Stores and News Dealers	11,562
		4522	Department Stores	15,835
		4523	General Merchandise Stores, including Warehouse Clubs and Supercenters	51,046
		4531	Florists	27,384
		4532	Office Supplies, Stationery, and Gift Stores	31,871
		4533	Used Merchandise Stores	33,762
		4539	Other Miscellaneous Store Retailers	78,477
		4543	Direct Selling Establishments	1,483
Pharmacies and gas stations	Essential	4461	Health and Personal Care Stores	160,088
		4471	Gasoline Stations	140,712
Offices	Essential	51	Information (media, software, and data)	72,241
		52	Finance and Insurance	423,536
		53	Real Estate Rental and Lending	687,357
		54	Professional, Scientific, and Technical Services	140,747
		55	Management of Companies and Enterprises	9,773
		561	Administrative and Support Services	21,170
		92	Public Administration	75,893
Education	Essential	6111	Elementary and Secondary Schools	126,556
		6112	Junior Colleges	2,716

Continued on next page

Table B.1 – continued from previous page

Category	Classification	NAICS	Description	Number of Venues
Education (cont.)	Essential	6113	Colleges, Universities, and Professional Schools	13,930
		6114	Business Schools and Computer and Management Training	351
		6115	Technical and Trade Schools	2,339
		6116	Other Schools and Instruction	60,023
		6117	Educational Support Services	247
Medical services	Essential	6211	Offices of Physicians	334,652
		6212	Offices of Dentists	154,454
		6213	Offices of Other Health Practitioners	187,244
		6214	Outpatient Care Centers	30,083
		6215	Medical and Diagnostic Laboratories	21,773
		6216	Home Health Care Services	26,671
		6219	Other Ambulatory Health Care Services	18,105
		6221	General Medical and Surgical Hospitals	21,841
		6222	Psychiatric and Substance Abuse Hospitals	477
		6223	Specialty Hospitals	15,864
Residential and elderly care	Essential	6231	Nursing Care Facilities (Skilled Nursing Facilities)	28,276
		6233	Continuing Care Retirement Communities and Assisted Living Facilities	27,305
		6241	Individual and Family Services	17,120
		6242	Community Food and Housing, and Emergency and Other Relief Services	3,592
		6244	Child Day Care Services	99,691
Recreation and entertainment	Non-essential	7111	Performing Arts Companies	1,028
		7112	Spectator Sports	4,959
		7113	Promoters of Performing Arts, Sports, and Similar Events	3,393
		7115	Independent Artists, Writers, and Performers	3
		7121	Museums, Historical Sites, and Similar Institutions	153,682
		7131	Amusement Parks and Arcades	7,116
		7132	Gambling Industries	2,747
		7139	Other Amusement and Recreation Industries	144,918
Hotels and lodging	Non-essential	7211	Traveler Accommodation	62,189
		7212	RV (Recreational Vehicle) Parks and Recreational Camps	8,605
Restaurants	Non-essential	7223	Special Food Services	8,485
		7224	Drinking Places (Alcoholic Beverages)	55,384
		7225	Restaurants and Other Eating Places	703,695
Other essential services	Essential	8111	Automotive Repair and Maintenance	234,736
		8112	Electronic and Precision Equipment Repair and Maintenance	7,565
		8114	Personal and Household Goods Repair and Maintenance	327,892
		8122	Death Care Services	154,890
		8123	Drycleaning and Laundry Services	14,172
		8129	Other Personal Services	51,692
		8131	Religious Organizations	393,328
		8132	Grantmaking and Giving Services	1,543
		8133	Social Advocacy Organizations	3,838
		8134	Civic and Social Organizations	39
		8139	Business, Professional, Labor, Political, and Similar Organizations	54
Other non-essential services	Non-essential	8121	Personal Care Services (e.g., barber shops, salons)	373,251
Total				6,786,255

C Appendix: Regression Estimates

This Appendix includes tables displaying additional regression estimates. Table C.1 displays estimates from the logistic regression model for the probability of death conditional on infection under the inclusion of geographic characteristics. Table C.2 displays OLS estimates for regressions of non-essential movement on age shares, mortality risk, and a variety of other controls and fixed effects.

Table C.1: Robustness of Fatality Rate Parameter Estimates: Demographics

Dependent Variable:	Death			
Model:	(1)	(2)	(3)	(4)
Variables				
(Intercept)	-7.385***	-3.579***	-8.997***	-2.547***
	(0.0181)	(0.0336)	(0.0408)	(0.0649)
Age = 0-17	-12.91	-12.90	-12.79	-12.81
	(16.20)	(15.93)	(15.96)	(15.85)
Age = 50-64	2.793***	2.693***	2.670***	2.596***
	(0.0191)	(0.0192)	(0.0193)	(0.0193)
Age = 65+	5.856***	5.737***	5.755***	5.626***
	(0.0179)	(0.0179)	(0.0180)	(0.0180)
Sex = Male	0.3686***	0.3559***	0.3494***	0.3478***
	(0.0050)	(0.0053)	(0.0053)	(0.0055)
Race = Asian	1.610***	0.7367***	0.5877***	0.2723***
	(0.0161)	(0.0177)	(0.0176)	(0.0190)
Race = Black	0.6139***	0.5650***	-0.1489***	-0.0491***
	(0.0078)	(0.0084)	(0.0090)	(0.0094)
Race = Hispanic	1.554***	1.042***	0.5734***	0.3495***
	(0.0084)	(0.0099)	(0.0106)	(0.0113)
Race = Multiple/Other	-0.4053***	-0.5759***	-1.122***	-1.013***
	(0.0357)	(0.0367)	(0.0404)	(0.0405)
Race = Native/PI	0.4700***	1.069***	0.7750***	0.9660***
	(0.0443)	(0.0461)	(0.0456)	(0.0465)
Percent Smoking		-3.643***		-10.23***
		(0.1102)		(0.1623)
Obesity Rate		-12.84***		-8.874***
		(0.0878)		(0.1008)
Physicians Per 100,000 Residents		-938.7***		-1,227.9***
		(11.06)		(12.77)
Percent Uninsured		-7.267***		-14.09***
		(0.0760)		(0.1075)

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Continued on next page

Table C.1 – *Continued from previous page*

Dependent Variable:			Death	
Model:	(1)	(2)	(3)	(4)
Particulate Matter 2.5		0.2607***		0.2100***
		(0.0016)		(0.0017)
Percent Rural			-1.578***	-1.401***
			(0.0221)	(0.0244)
Median Household Income			$1.5 \times 10^{-5***}$	$-1.26 \times 10^{-5***}$
			(3.12×10^{-7})	(4.04×10^{-7})
Percent Below Poverty Line			7.010***	6.140***
			(0.1306)	(0.1381)
Percent Without Vehicle			5.837***	3.261***
			(0.0478)	(0.0525)
Percent Without Computer Access			-3.191***	1.167***
			(0.0968)	(0.1108)
Percent Limited English			6.488***	6.400***
			(0.0727)	(0.1007)
Fit statistics				
Observations	4,660,804	4,660,804	4,660,804	4,660,804
Squared Correlation	0.20986	0.27838	0.28944	0.32933
Pseudo R^2	0.38949	0.44049	0.45077	0.47830
BIC	1,058,231.4	969,910.8	952,118.0	904,475.0

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*
IID standard-errors in parentheses

Notes: This table displays point estimates and standard errors from logistic regression models estimated using the CDC COVID-19 Case Surveillance Public Use Data with Geography, which has information on infected patients' demographics and health outcomes, and county-level geographic characteristics from the American Community Survey (ACS) and County Health Rankings (CHR) Analytic data. The outcome variable is an indicator for whether a patient died as a result of COVID-19 infection. The excluded categories (baseline characteristics) are age 18-49, sex female, and race white. Observations with missing values for the outcome variable or any of the covariates are excluded. Column 1 includes only demographic characteristics, column 2 includes geographic characteristics from the CHR data, column 3 includes geographic characteristics from the ACS data, and column 4 includes both sets of geographic characteristics.

Table C.2: Robustness of Movement Parameter Estimates: Age and Mortality Risk

Dependent Variable:	Non-Essential Visits				
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Share 10-19 Years	1.430***	1.334***	3.588***	2.316***	0.8216***
	(0.1238)	(0.0181)	(0.0250)	(0.0316)	(0.0246)
Share 20-29 Years	0.8908***	2.229***	3.480***	2.695***	1.926***
	(0.1209)	(0.0133)	(0.0178)	(0.0238)	(0.0184)
Share 30-39 Years	0.7519***	2.080***	3.647***	2.828***	1.937***
	(0.0932)	(0.0167)	(0.0226)	(0.0271)	(0.0210)
Share 40-49 Years	1.228***	1.136***	2.918***	2.377***	0.7755***
	(0.0852)	(0.0227)	(0.0301)	(0.0312)	(0.0243)
Share 50-59 Years	1.587***	0.8058***	2.863***	2.682***	0.5294***
	(0.0811)	(0.0202)	(0.0273)	(0.0277)	(0.0216)
Share 60-69 Years	1.226***	1.114***	2.900***	2.549***	0.7483***
	(0.0870)	(0.0222)	(0.0292)	(0.0296)	(0.0231)
Share 70-79 Years	1.613***	1.570***	3.493***	3.123***	1.120***
	(0.1180)	(0.0296)	(0.0375)	(0.0381)	(0.0296)
Share 80+ Years	0.6787***	1.070***	3.123***	2.733***	0.5401***
	(0.0998)	(0.0283)	(0.0366)	(0.0390)	(0.0303)
Mortality Risk × Share 10-19 Years	2.523	-5.575***	-2.667	-4.480*	1.227
	(5.374)	(1.943)	(2.457)	(2.441)	(1.889)
Mortality Risk × Share 20-29 Years	-2.302	-11.30***	-8.454***	-9.016***	-12.61***
	(1.656)	(0.5126)	(0.6478)	(0.6434)	(0.4981)
Mortality Risk × Share 30-39 Years	1.621	-2.609***	-2.842***	-2.702***	-2.879***
	(0.9853)	(0.2414)	(0.3051)	(0.3030)	(0.2346)
Mortality Risk × Share 40-49 Years	0.4290	0.1106	-0.7865***	-0.7793***	0.2194*
	(0.2967)	(0.1259)	(0.1592)	(0.1581)	(0.1224)
Mortality Risk × Share 50-59 Years	-0.0740	0.0063	-0.2372***	-0.2582***	0.0532
	(0.0811)	(0.0437)	(0.0553)	(0.0549)	(0.0425)
Mortality Risk × Share 60-69 Years	-0.1190***	-0.0631***	-0.2048***	-0.1996***	-0.0444***
	(0.0255)	(0.0172)	(0.0217)	(0.0216)	(0.0167)
Mortality Risk × Share 70-79 Years	-0.1149***	-0.0923***	-0.1832***	-0.1808***	-0.1032***
	(0.0144)	(0.0097)	(0.0122)	(0.0121)	(0.0094)
Mortality Risk × Share 80+ Years	-0.0535***	-0.0395***	-0.1121***	-0.1101***	-0.0457***
	(0.0066)	(0.0048)	(0.0061)	(0.0060)	(0.0047)
Essential Visits		0.7173***			0.7225***
		(0.0011)			(0.0011)
Share Essential Workers			0.3277***	0.2835***	-0.0521***
			(0.0116)	(0.0141)	(0.0109)
Median Household Income			7.51×10^{-6}***	3.26×10^{-6}***	1.48×10^{-6}***
			(5.21×10^{-8})	(7.95×10^{-8})	(6.16×10^{-8})
Share without Health Insurance			-0.4475***	-0.4642***	-0.2123***
			(0.0283)	(0.0285)	(0.0220)
Percent without Internet			-0.2581***	-0.2413***	-0.3804***
			(0.0161)	(0.0187)	(0.0144)
Share Below Poverty Level				0.1523***	0.2705***
				(0.0165)	(0.0128)
Share with Bachelor's Degree				0.7404***	0.6638***

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Continued on next page

Table C.2 – *Continued from previous page*

Dependent Variable:	Non-Essential Visits				
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
				(0.0132)	(0.0102)
Share Renters				-0.1638***	-0.3569***
				(0.0097)	(0.0075)
Average Household Size				0.2547***	-0.0195***
				(0.0038)	(0.0030)
Percent Crowding				-0.0713	0.2920***
				(0.0442)	(0.0342)
Percent Mobile Home				-0.0266	-0.0639***
				(0.0308)	(0.0238)
Percent without Phone				0.1358**	-0.2937***
				(0.0528)	(0.0409)
Percent without Vehicle				0.1222***	0.9934***
				(0.0140)	(0.0109)
Percent without Plumbing				-0.7821***	-0.0294
				(0.0467)	(0.0362)
Percent without Kitchen				-0.1619***	-0.0419
				(0.0423)	(0.0327)
Percent Group Quarters				0.0742***	0.5087***
				(0.0233)	(0.0180)
Fixed-effects					
Week	Yes				
County	Yes				
Fit statistics					
Observations	630,801	630,801	630,801	630,801	630,801
R^2	0.28061	0.41855	0.07121	0.08416	0.45123
Within R^2	0.00936				

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: This table displays ordinary least squares (OLS) estimates of the parameters relating age and mortality risk to non-essential visits. The dependent variable is the number of non-essential visits per cell phone tracked in a block group in a given week. Non-essential visits are measured using SafeGraph data and classified based on NAICS code as described in Appendix B. Age shares and other block group characteristics are derived from American Community Survey 2015-2019 5-year estimates. The mortality risk measure is constructed as described in Appendix A. In columns two through five, the coefficients on the age shares estimate the baseline level of non-essential visits of that age group; there is no excluded group or constant. In column one, there are county and week fixed effects, so the same interpretation does not hold. In all columns, the coefficients on the age shares interacted with mortality risk estimate the responsiveness a person of that age to mortality risk.