Essays on Health Care and Inequality

HARVARD UNIVERSITY
Graduate School of Arts and Sciences

DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the

Department of Public Policy

have examined a dissertation entitled

"Essays on Health Care and Inequality"

presented by Frina Lin

candidate for the degree of Doctor of Philosophy and hereby certify that it is worthy of acceptance.

Signature ____________________

Typed name: Professor Marcella Alsan, Chair

Signature ____________________

Typed name: Professor David Cutler

Signature ____________________

Typed name: Professor Will Dobbie

Date: August 22, 2022

Essays on Health Care and Inequality

A dissertation presented

by

Frina Lin

to

The Department of Public Policy

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Public Policy

Harvard University

Cambridge, Massachusetts

August 2022

Dissertation Advisor: Professor Marcella Alsan Frina Lin

Essays on Health Care and Inequality

Abstract

This dissertation consists of three chapters analyzing the role of institutional factors in observed health inequities.

In the first chapter, I measure how differences in neighborhood environments contribute to Black-white disparities in primary care outcomes. I study this question in the context of preventive health care take-up among Medicare enrollees, a setting with full insurance and yet persistent racial gaps in recommended care. Leveraging variation from Medicare enrollees who move across areas, I show that local environments have sizable and immediate impacts on the probability of receiving recommended care and that Black enrollees on average live in areas with lower causal effects on outcomes. Within areas, Black enrollees receive worse care than white enrollees when residential racial segregation is greater. Using this insight, I estimate that residential sorting across Primary Care Service Areas (PCSAs) accounts for 28% of the Black-white gap in having an annual primary care visit, while heterogeneous effects by race within PCSAs account for 16% of the gap. Although individual-level differences such as preferences for health care utilization contribute to racial disparities in outcomes considered, local area effects play a substantial role as well.

In the second chapter, I study whether heterogeneous hospital quality generates different local health care environments by race in the context of emergency care. I test whether there are differences in the ordered quality ranking of local hospitals for Black and white emergency patients. To control for underlying factors that affect hospital choice, I exploit ambulance company preferences as an instrument for hospital characteristics and estimate the impact of being treated at a hospital with a higher share of patients who are Black.

Using mortality following hospitalization as the primary outcome, I find substantial race-specific effects, with Black patients experiencing better outcomes at hospitals with greater Black patient shares in the medium to long run. The results are consistent with a hospital choice model that exhibits institutional comparative advantage and positive Roy selection by patient race, pointing to the limitations of broad-based hospital quality measures.

In the third chapter, I examine trends in the trajectories of Black medical school applicants over the years 1979-2020. I find that Black applicants have grown to 10% of the applicant pool in recent years, but acceptance rates remain below those of white and Asian applicants. For Black students who do matriculate to medical school, graduation rates lag behind those of white and Asian medical students. Among the cohorts of graduated MDs, medical schools affiliated with historically Black colleges and universities (HBCUs) continue to play an important role in graduating Black physicians, accounting for 14.9% of Black physicians in 1984-1999 and 14.7% in 2000-2015. Taken together, the evidence highlights continuing gaps for Black students in the physician pipeline, with need for targeted actions.

Contents

List of Tables

List of Figures

Chapter 1

Neighborhoods and Racial Disparities in Preventive Care

Abstract

To what extent do unequal environments contribute to racial disparities in outcomes? I study this question in the context of preventive health care take-up among Medicare enrollees, a setting with full insurance and yet persistent racial gaps in recommended care. Leveraging variation from Medicare enrollees who move across areas, I show that local environments have sizable and immediate impacts on the probability of receiving recommended care and that Black enrollees on average live in areas with lower causal effects on outcomes. Within areas, Black enrollees receive worse care than white enrollees when residential racial segregation is greater. Using this insight, I estimate that residential sorting across Primary Care Service Areas (PCSAs) accounts for 28% of the Black-white gap in having an annual primary care visit, while heterogeneous effects by race within PCSAs account for 16% of the gap. Although individual-level differences such as preferences for health care utilization contribute to racial disparities in outcomes considered, local area effects play a substantial role as well.

1.1 Introduction

Neighborhoods are an important unit of policy intervention and shape outcomes across health, education, work and income. Across many of these domains, Black and other minority families are more likely to live in areas with observably worse characteristics, such as higher mortality rates, lower public spending on education, and higher poverty rates. However, the extent to which worse neighborhood environments contribute causally to racial disparities in outcomes is unknown. Variation in observed neighborhood characteristics could reflect selection on individual characteristics, as well as simultaneously determined neighborhood effects on residents' outcomes. In this paper, I overcome individual selection problems using quasi-experimental variation from individuals who move across neighborhoods. Examining longitudinal data from a sample of Medicare enrollees, I estimate the causal effect of neighborhood environments on racial disparities for a widely recommended and observable outcome: presence of an annual visit to a primary care provider (PCP).

Racial disparities in health, in particular Black-white disparities, are a persistent feature of U.S. history (Institute of Medicine, 2003). Black men live on average 4.4 years fewer than white men (Arias and Xu, 2015), and Black-white disparities have been documented across many facets of health care including access to routine care (Gray *et al.*, 2017), referrals for diagnostic imaging (Colwell *et al.*, 2022), procedure rates for transplants and surgeries (Kaiser Family Foundation, 2002; Malek *et al.*, 2011), medication prescription rates (Schore *et al.*, 2004), and hospitalizations for uncontrolled chronic conditions (Gray *et al.*, 2017). In particular, differences in primary and preventive care utilization are thought to explain a large portion of the Black-white gap in life expectancy (Harper *et al.*, 2012).

Recent literature has shown the importance of area factors in addition to individual factors in determining important outcomes. Fundamentally, the counties in which children grow up shape earnings and college attendance rates, in proportion to the amount of childhood time spent in an area (Chetty and Hendren, 2017a,b). In the realm of health care utilization, 50-60 percent of geographic variation in overall utilization is attributable to supply factors at the Hospital Referral Region level (Finkelstein *et al.*, 2016), and 30 percent of geographic variation in opioid abuse is attributable to place-specific factors at the county level (Finkelstein *et al.*, 2018). Mortality is also immediately and substantially affected by residental location: individuals who move to cities and commuting zones with higher observed life expectancy among non-movers see reductions in post-move mortality (Finkelstein *et al.*, 2019; Deryugina and Molitor, 2020).

In this paper, I build on existing literature, using identification from individuals who move across areas, referred to as “movers," to control for unobserved individual characteristics and to isolate the causal effect of local environments on Black-white disparities in primary care take-up. The empirical strategy captures the immediate effect of moving a person to a different neighborhood by observing changes in primary care utilization when a patient moves across neighborhoods. I define neighborhoods to be collections of zip codes called “Primary Care Service Areas" (PCSAs), defined by the Dartmouth Atlas of Health Care to approximate areas within which most residents seek primary care.

In particular, I conceptualize two ways in which neighborhood environments may contribute causally to racial disparities in outcomes. First, Black and white populations in the U.S. could live in different neighborhoods, such that the neighborhoods where white families live have better causal effects on outcomes on average than the neighborhoods where Black families live. Neighborhoods with stronger causal effects could have greater access to primary care physicians and higher quality primary care. I call this channel, driven by racial sorting across neighborhoods, *sorting*.

Second, within the same neighborhood, there could still be differences in effectiveness of primary care by race. Using PCSAs to delineate neighborhoods, I posit that geographic access to primary care providers and neighborhood amenities would be largely constant within each PCSA. Although residential segregation within neighborhoods remains possible, the scope for differences in residential location to directly cause differences in primary care effectiveness by race is limited within a neighborhood. On the other hand, I consider the possibility that beneficiaries of different race could experience the same health care environment differently. I call this channel, driven by within-neighborhood race-specific differences in costs of receiving care, *efficacy*. I argue that it is important to consider and quantify both the sorting and efficacy channels to fully account for the effect of local neighborhood environments on racial disparities.

I estimate the role of neighborhood sorting and efficacy channels on the Black-white gap in primary care utilization using data from a 20% sample of Medicare fee-for-service enrollees from 1999 to 2017. I focus on the presence of an annual primary care provider (PCP) visit as the main outcome variable, as this is a widely recommended preventive care measure that is observable in claims data. In the raw data, I document a substantial Black-white gap in the presence of an annual primary care visit, which persists over 1999-2017 and ranges from 8-12pp.

Building on the model in Finkelstein *et al.* (2016), I introduce a model of health care demand and supply which implies that patients' choice of primary care utilization can be written as a combination of patient fixed effects, location by race fixed effects, and a vector of time-varying controls. The specification allows for movers to have primary care utilization levels which are systematically different and correlated with their origin and destination locations. It also allows for movers to have trends in utilization around the time of move. I conduct the analysis separately by patient race. They key identifying assumptions are that there are no differential trends in utilization which are correlated with migrants' origin and destination locations around the time of move, for movers of each race.

The estimating equation is highly related to a set of papers, which use movers designs and two-way fixed effects models to control for individual variation and selection. In labor economics, Abowd *et al.* (1999) (AKM) and others (e.g. Card *et al.* (2013, 2016); Song *et al.* (2019)) have employed two-way fixed effects models with longitudinal data on workers who switch firms to separate person-specific and firm-specific drivers of wage variation. In health economics, Finkelstein *et al.* (2016) and others (e.g. Finkelstein *et al.* (2018); Molitor (2018); Allcott *et al.* (2019); Finkelstein *et al.* (2019); Deryugina and Molitor (2020)) have used movers designs to separate person-specific and place-specific factors driving variation in health care utilization and mortality. Recent literature has questioned the sensitivity of and assumptions underlying analyses employing the two-way fixed effect model and estimation. In particular, these critiques point out biases that arise in the variances of the estimated fixed effects due to limited mobility and weak identification with many regressors, and propose bias-correction methods (Bonhomme *et al.*, 2020; Kline *et al.*, 2020). Where applicable I implement the bias correction of Kline *et al.* (2020), adapting the leave-out estimator methodology for the patient-place fixed effects model.

Following the framework of Card *et al.* (2016), who decompose firm-specific drivers of the gender wage gap into sorting and bargaining channels, I estimate a two-way fixed effects model and measure sorting and efficacy channels via an Oaxaca-style decomposition (Oaxaca, 1973) of race-specific neighborhood effects on primary care utilization. A key issue for assessing the contribution of the efficacy channel (i.e. the "bargaining" channel in Card *et al.* (2016)) is the need to define relevant reference groups for each race (Oaxaca and Ransom, 1999). I define a normalization based on within-neighborhood measures of residential racial segregation to provide a lower bound estimate of differential efficacy for Black enrollees.

As in AKM and Card *et al.* (2016), I begin the empirical analysis with descriptive evidence on the presence of neighborhood effects on primary care take-up and the plausibility of the exogenous mobility assumptions needed to measure them via the two-way fixed effects model with ordinary least squares (OLS) estimation. I find that patterns of primary care take-up before and after a move are consistent with the indentifying assumptions for both Black and white Medicare enrollees.

I then estimate separate AKM models by enrollee race. I find substantial variation in neighborhood-specific causal effects, with neighborhoods one standard deviation higher in primary care outcomes having about a 10pp greater impact on having an annual PCP visit for both white and Black enrollees. The bias-corrected estimates of the standard deviation of person-specific and place-specific effects by race show that person effects vary more widely than place effects for enrollees of both races, and place effects vary similary for Black and white enrollees. I also find greater variation in person effects for Black enrollees than for white enrollees. This result is consistent with literature highlighting the particular role of factors such as trust in health care on health care utilization of Black and minority patients.

Turning to estimates of the contribution and decomposition of neighborhood effects on the Black-white gap in having an annual PCP visit, I estimate that the total contribution of neighborhood effects to the Black-white gap is 44%, with 28% of the gap attributable to the sorting channel and 16% of the gap attributable to the efficacy channel. These results indicate that accounting only for the sorting channel would understate the role of neighborhood environments on the Black-white gap. In addition, I find that the portion of the Black-white gap in having an annual PCP visit attributable to neighborhood effects is consistent around 40-50 percent regardless of age among the sample of age 65+ Medicare enrollees, with similar results for each 5-year age bin observed in the data.

Finally, I examine the results of the model varying the geographic unit considered to be the "local area". I find that, again, place effects consistently drive 40-45 percent of the Black-white gap in having an annual PCP visit, through a combination of sorting and efficacy channels. At all local area levels, measures of residential racial segregation systematically predict areas where Black and white causal effects diverge. However, I find that when local areas are defined as the larger Hospital Referral Regions (HRRs), the contribution of the sorting channel drops to zero and all place effects are realized through Black-white differences in within-HRR efficacy. This result indicates that regional residential sorting is not driving the Black-white gap but that micro-level residential segregation within regions consistently predicts Black enrollees being left behind.

This paper contributes to several strands of literature. First, a long literature has documented the scope of health inequalities, including how health-related outcomes differ widely across individual characteristics such as income and education (e.g. Cutler and Lleras-Muney (2010); Chetty *et al.* (2016)), which play a role in racial disparities in health (e.g. Geruso (2012); Lahiri and Pulungan (2021)). However, recent literature has also suggested important roles for context, social structure, and neighborhood amenities in determining health inequities, especially by race (e.g. Williams and Sternthal (2010); Adler *et al.* (2016)). This paper empirically quantifies the role of immediate neighborhood context versus constant individual characteristics. The results are consistent with existing literature: both individual and neighborhood factors are found to play substantial roles in driving the Black-white gap in primary care utilization. Indeed, although individual characteristics drive the majority of the Black-white gap, neighborhood effects, including striking within-neighborhood differences in efficacy, cannot be overlooked.

Second, this paper is consistent with and adds to literature measuring causal channels through which Black and white patients experience different health care environments which lead to differing health outcomes. In particular, reduced trust in health care and ease of communication with health care providers among Black men have been found to impact primary care utilization and mortality outcomes (Alsan and Wanamaker, 2018; Alsan *et al.*, 2019). This paper builds on estimates of specific causal mechanisms, conceptualizing the broader net impact of local environments on racial disparities, through *sorting* and *efficacy* channels. Importantly, I find new evidence for the efficacy channel, which encourages additional research into ways in which within-neighborhood environments differ for Black and white patients.

Finally, I add to literature connecting residential segregation to worse health outcomes for minorities. Residential segregation has been found to correlate with Black patients receiving care at lower quality hospitals than their white peers, even though Black patients on average live closer to better quality hospitals (Dimick *et al.*, 2013). Although it remains unclear the mechanisms through which racial segregation predicts worse outcomes for Black individuals, this paper suggests that greater focus on racial segregation, especially micro-segregation within small geographic areas, is warranted.

1.2 Data

1.2.1 Medicare Data

My primary data source is a 20 percent sample of Medicare fee-for-service beneficiaries. For this set of individuals, I observe health care claims across settings billed to Medicare between 1999 and 2017, along with annual enrollment information and demographic characteristics. I restrict to beneficiary-year periods in which the beneficiary is aged 65 to 99. This follows prior literature and ensures that all beneficiaries are age-eligible for Medicare, as opposed to being eligible through disability. As of 2020, 62.8 million individuals were enrolled in Medicare, of which 60% were enrolled in fee-for-service, Original Medicare and 40% in Medicare Advantage plans (CMS, 2021).

I construct the outcome variable of having an annual visit to a primary care provider (PCP) by tagging claims for health care services within a set of provider specialties and visit location types. Following Goodman *et al.* (2010), I code PCP visits as claims billed from a clinician with a specialty of general practice, family practice, internal medicine, nurse practitioner, physician assistant, or clinic and located in an office, public health clinic, rural health clinic, federally qualified health center, walk-in retail health clinic, or independent clinic. This definition captures a range of relevant primary care visits.

For my analysis, I restrict to beneficiary-year observations in which the beneficiary is enrolled in Medicare Parts A and B with no Medicare Advantage enrollment. Medicare Part A covers fee-for-service inpatient care in hospitals, as well as skilled nursing facility, hospice, and home health care. Medicare Part B covers services from physicians and other health care providers, outpatient care, durable medical equipment, and preventive services. This restriction removes enrollee-years for which claims for services are not observed.

I observe additional demographic variables at the beneficiary-year level, including age, gender, race, and zip code of residence. The race variable comes from a linkage to self-reported race from the Social Security Administration. For the purpose of this study, I restrict to beneficiaries whose race is Black or white.

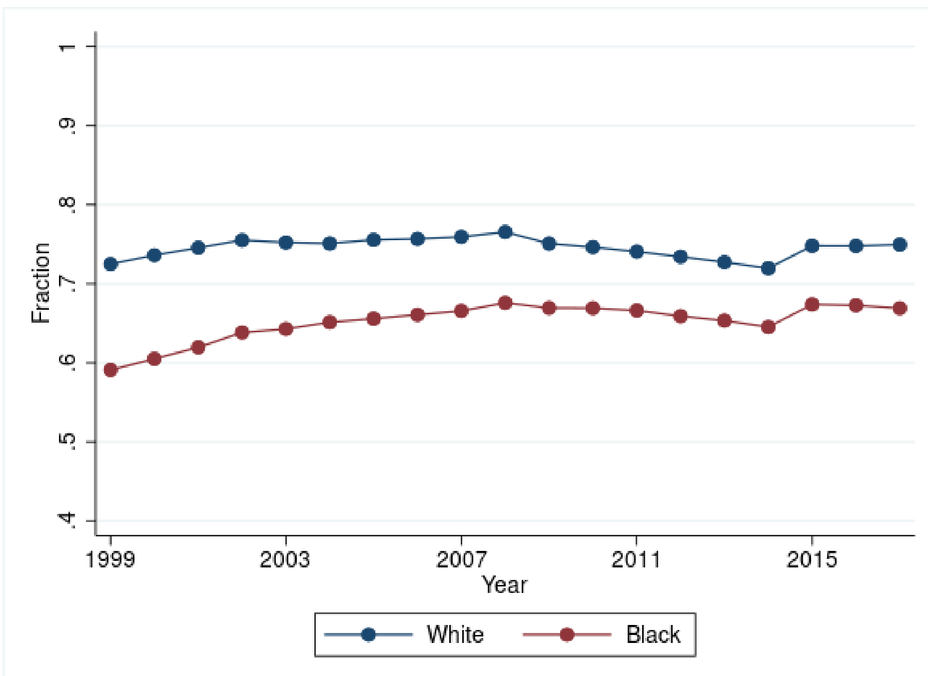
Zip code of residence is based on the address on file for cash benefits to the beneficiary, including Social Security, as of December 31 of each year. I label enrollees as moving in a given year if their zip code on file changes from the zip code observed the prior year. Using this information on zip codes, I construct two mutually exclusive samples of individuals. The first is a *stayer* sample of enrollees whose zip code never changes for the years in which they are observed in the data. The second is a *mover* sample of enrollees who move residence during the sample period.

1.2.2 Stayer Sample

To contruct the stayer sample, I include beneficiaries whose zip code of residence does not change across all enrolled years over 1999-2017. To address concerns about differences in age distribution across years and areas, which may contribute to variation in PCP visit rates, I restrict to stayers between the ages of 65 and 75. This results in 31 million white enrollee-year observations and 3 million Black enrollee-year observations, which I use to compute descriptive statistics on the rate of PCP visits, across years and geographic areas by race.

Figure 1.1 shows that over the sample period of 1999 to 2017 there has been a persistent Black-white difference in presence of an annual PCP visit. The gap narrowed from a 12.5 percentage point difference in years 2000-2002 to a 7.5 percentage point difference in years 2014-2016 but remains substantial in the most recent periods. Given these descriptive statistics, this paper focuses on the approximately 10pp gap in annual PCP visit rates among Medicare enrollees, a question which is empirically relevant. The presence of a sizable gap in annual PCP visit rates between Black and white Medicare beneficiaries, who have equivalent health insurance coverage, motivates research into causal factors driving the difference.

Figure 1.1: *Racial Disparities in Primary Care Visits Among Medicare Enrollees*



Notes: Figure shows fraction of Medicare beneficiaries with any primary care visit within each calendar year. Computed from a 20 percent sample of Traditional Medicare (fee-for-service) enrollees, aged 65-75, who have no changes in zip code of residence over 1999-2017.

1.2.3 Mover Sample

To construct the mover sample, I include beneficiaries whose zip code of residence changes once within the sample period of 1999-2017. The empirical strategy relies on enrollees whose external health care environment changes at the time of a move, so I define the sample to include only beneficiaries who move to a new Hospital Referral Region (HRR). The 306 HRRs approximate broad market areas for referrals to specialty hospital care[1]. Empirically, patients receive a large majority of health care services within their HRR (Finkelstein *et al.*, 2016).

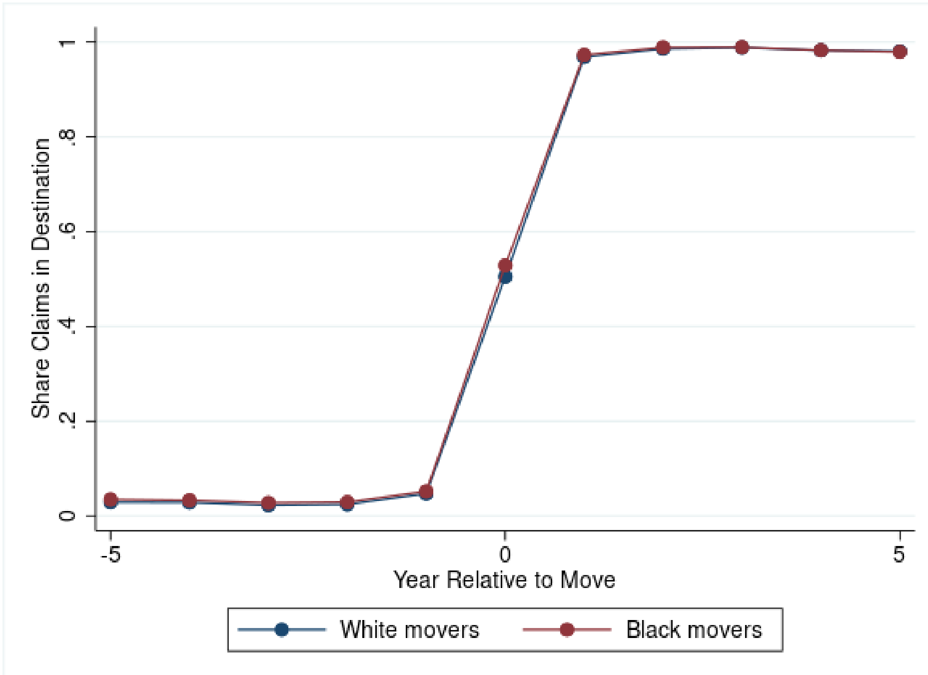
Following Finkelstein *et al.* (2016), I further restrict the mover sample to beneficiaries whose health care claims in the destination HRR as a share of claims in either the destination or origin HRR changes by more than 0.75 when comparing the five years before and after their zip code change. This restriction ensures that beneficiaries in the mover sample have actually moved and transitioned their health care services in accordance with the move. Without the restriction, beneficiaries identified as movers could have changed their mailing address on file without changing their residence, for example if they decide to have their Social Security checks sent to a child who is handling their finances, or if they have multiple residences both before and after the move.

I denote the year in which a beneficiary's zip code changes as relative year 0 and include beneficiary-year observations in the analysis data within the five years before and after the move, from relative year -5 to 5. Figure 1.2 shows that enrollees in the final sample sharply move their claims locations around the year of the move, as we would expect. The pattern is similar for Black and white enrollees in the mover sample.

Table 1.1 shows summary statistics for the mover sample, by enrollee race. The Black mover sample is similar to the white mover sample in age and gender. Black movers are observed in the data for slightly fewer years, with 5.41 years observed on average within the 11 years surrounding and including the move year, in comparison to 6.27 years observed

[1]HRRs are defined by the Dartmouth Atlas of Health Care to include at least one hospital that performs major cardiovascular procedures and have a population of at least 120,000. Each HRR is made up of zip codes for which the highest proportion of cardiovascular procedures are referred to a hospital within its boundaries.

Figure 1.2: *Share of Claims in Destination by Relative Year*



Notes: Figure shows the share of movers' health care claims located in their destination HRR, among those located in either their origin or destination HRR, in the five years before and after the year in which their zip code changes. Shares are computed from 2,414,308 patient-years for the white mover series and 123,520 patient-years for the Black mover series.

Table 1.1: *Summary of Movers Sample*

	White	Black
PCP visit	0.807	0.726
Female	0.599	0.635
Age	73.8	73.0
Years Observed	6.27	5.41
# of Movers	342,586	20,170
# of Observations	2,148,720	109,193

Notes: Table shows summary statistics for the set of beneficiary by year observations for the sample of Medicare enrollees who moved exactly once within 1999 to 2017.

for white movers. The Black-white gap in presence of an annual PCP visit is 8.1pp, which is similar to the gap in PCP visit rates observed in the stayer sample. In total, the mover sample for analysis includes 109,193 observations of 20,170 unique Black beneficiary movers and 2,148,720 observations across 342,586 unique white beneficiary movers.

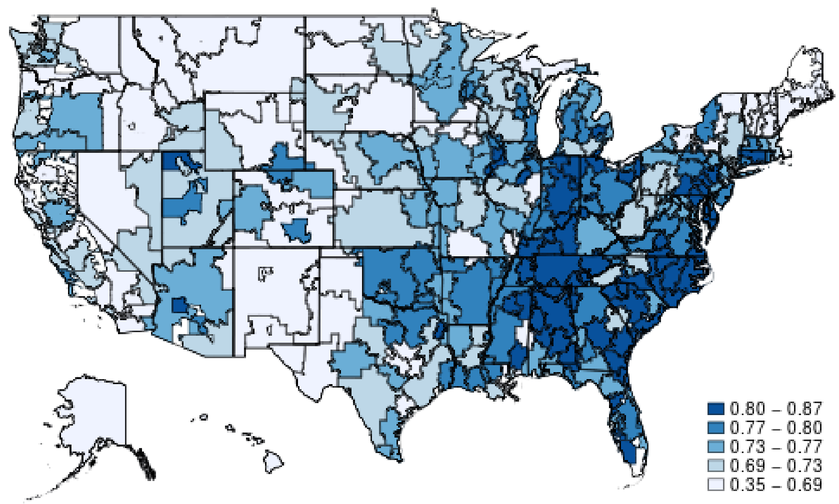
1.2.4 Neighborhood Characteristics

I define neighborhoods to be Primary Care Service Areas (PCSAs). The 6,542 PCSAs are defined by the Dartmouth Atlas of Health Care and constructed to approximate markets for primary care. Each PCSA was delineated using Medicare claims data to include a zip code with one or more primary care providers, along with additional contiguous zip codes whose populations seek a plurality of their primary care from the same providers. The median PCSA contains 3 zip codes and a population of 15,000. Because we want to measure the effects of local health care environments on primary care, PCSAs are a natural geographic unit to ground the analysis on.

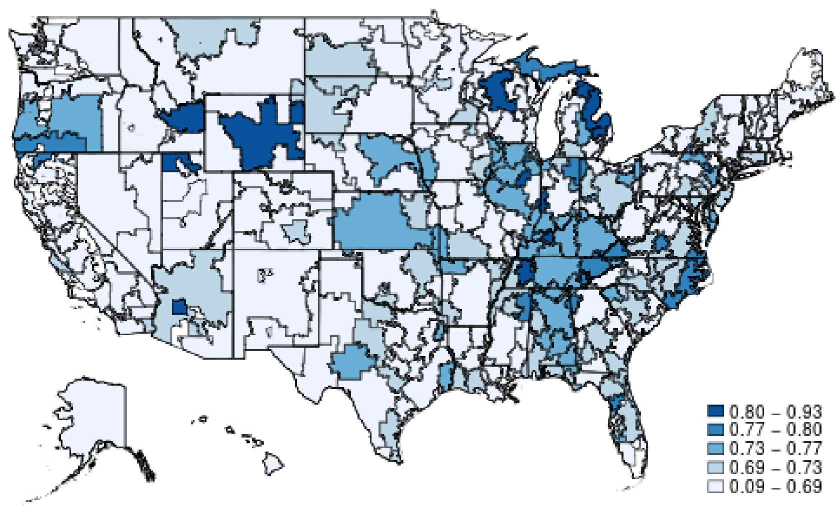
First, at the larger (regional) HRR level, Figure 1.3 shows that there is substantial variation in rates of annual PCP visits across local areas and by race, using data from the stayer sample. As noted by prior literature on geographic variation in health care, this broad variation could be due to *individual factors* or *area factors*. On one hand, patient characteristics such as health status and preferences could differ across areas, generating variation in observed take-up of annual PCP visits. On the other hand, patients could be similar across areas in their underlying likelihood of seeking care, and place-specific variables such as doctors' incentives and beliefs and the health care market structure could drive the observed variation. The scope of the geographic variation, ranging from from 69% of patients receiving annual PCP visits in the 20th percentile Hospital Referral Region (HRR) to 80% receiving annual PCP visits in the 80th percentile HRR, is of a similar order of magnitude to the observed Black-white gap.

Panel B of Figure 1.3 shows that Black enrollees have lower rates of annual PCP visits in almost all areas. However, there remains substantial variation across areas, including some

Figure 1.3: *Geographic Variation in Primary Care Visits*



(a) *White*



(b) *Black*

Notes: Figure shows fraction of Medicare beneficiaries, by Hospital Referral Region (HRR) of residence, with any primary care visit within the calendar year, averaged over years 1999 to 2017. Computed from a 20 percent sample of Traditional Medicare (fee-for-service) enrollees, aged 65-75, who have no changes in zip code of residence over the sample period.

areas with relatively high rates of annual PCP visits.

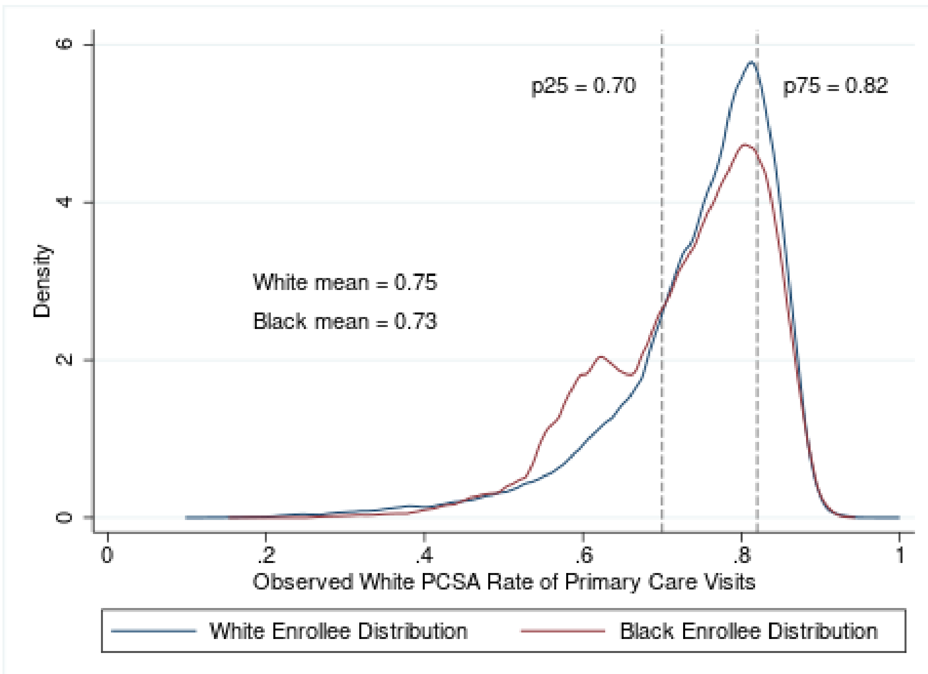
Similarly, there is wide variation in annual PCP visit rates across neighborhoods (PCSAs) as shown in the blue and red lines in Figure 1.4. For the combined Black and white mover samples, the 25th percentile population-weighted neighborhood has a neighborhood annual PCP visit rate among White Medicare stayers of 70%, and the 75th percentile neighborhood has an annual PCP visit rate of 82%. Again, this 25th to 75th percentile difference in neighborhood characteristics is of a similar order of magnitude to the observed Black-white gap, suggesting that neighborhood effects could drive part of the disparity.

Indeed, comparing the red series to the blue series in Figure 1.4, we see that enrollees in the Black mover sample have higher density in neighborhoods with lower rates of annual PCP visits, so that the mean neighborhood annual PCP visit rate for Black movers is 73% compared to a mean neighborhood annual PCP visit rate for white movers of 75%. The difference in distribution of Black and white movers across PCSAs suggests that, if we assume neighborhood rates entirely reflect causal effects of neighborhoods, the difference in sorting across neighborhoods would explain 2pp or 25% of the observed Black-white gap. This 25% can serve as a benchmark of the naive estimate of the role of neighborhood effects on the Black-white gap in the movers sample.

Finally, I analyze other neighborhood characteristics including a measure of residential racial segregation within neighborhoods. I use data from Chetty *et al.* (2016), which reports a Theil index of racial segregation constructed within counties and across census tracts accounting for four racial groups: white, Black, Hispanic, and other[2]. I interpolate these data to the neighborhood level, to have a measure of racial differences even within the primary care neighborhoods defined as the main unit of analysis for local health care environments.

[2]See Chetty *et al.* (2014) for more information about construction of the Theil index.

Figure 1.4: *Distribution of Medicare Enrollees Across PCSAs*



Notes: Figure shows smoothed distribution of the area-level rate of primary care visits in the Primary Care Service Areas (PCSAs) where white Medicare enrollees live and in the PCSAs where Black Medicare enrollees live. Area-level rates of primary care visits are constructed at the PCSA level using data from white Medicare enrollees, aged 65-75, who have no changes in zip code of residence over the sample period.

1.3 Empirical Framework

1.3.1 Model

I present a model of supply and demand for preventive care services with neighborhood (i.e. PCSA) effects that may differ by race. I consider a population of patients i in year t utilizing preventive care $y_{it} \in R^{+}$. Utilization depends on individual time-varying health status h_{it} as well as individual time-constant preferences η_i. Enrollees also face cost $c_{rjt}(y)$ of receiving care, given race $r \in \{white, black\}$ and neighborhood j in year t. Examples of these costs, which may differ by neighborhood and patient race, include travel costs, patient co-pays, and difficulties scheduling an appointment, as well as frictions in communicating with or discomfort visiting a PCP.

Adapting the patients' utility function in Finkelstein *et al.* (2016), we write that enrollee i, of race r living in neighborhood j in year t, faces the following expected continuation utility:

$$u(y|h_{it}, \eta_i, r, j, t) = -\frac{1}{2}(y - h_{it})^2 + \eta_i y - c_{rjt}(y). \tag{1.1}$$

Assuming that $c_{rjt}(y)$ is linear in y, we have that expected utility is maximized at:

$$y^{*}_{itrj} = h_{it} + \eta_i - c'_{rjt}(). \tag{1.2}$$

In the data on Medicare enrollees, we observe choices of y_{it} given $r(i)$ and $j(i,t)$ across neighborhoods and time. We are interested in identifying attributes of the marginal cost parameter $c'_{rjt}()$, including its variance across neighborhoods and means by race. Assuming that the expectation of y^{*}_{itrj} given the data observed depends only on a patient fixed effect and a vector of observables x_{it} and that marginal costs of receiving care are additively separable in j and t, we can map Equation 1.2 onto the following estimating equation:

$$y_{it} = \alpha_i + \gamma^{r(i)}_{j(i,t)} + x'_{it}\beta + \tau_t + \epsilon_{it}. \tag{1.3}$$

Equation 1.3 is a standard two-way fixed effects model with enrollee fixed effects α_i and neighborhood by race fixed effects $\gamma^{r(i)}_{j(i,t)}$, in addition to coefficients β for observed time-varying enrollee characteristics and year fixed effects τ_t.

In the specification for analysis, the utilization quantity y will be a binary variable indicating whether an enrollee had any PCP visit during the year and x_{it} will include relative year fixed effects $\rho_{k(i,t)}$, where the relative year is $k(i,t) = t - t_i^*$ for a mover who moves during year t_i^*. Including the relative year fixed effects allows for the possiblity that the timing of a move is correlated with shocks to health status.

1.3.2 Identification

I estimate models based on Equation 1.3 using OLS, generating race-specific neighborhood causal effects $\hat{\gamma}_j^r$. For these estimates to be unbiased, the underlying data must exhibit exogenous mobility and additive neighborhood effects.

The exogenous mobility assumption could be violated if enrollees who are becoming relatively more health-conscious or have received adverse health shocks prior to a move systematically relocate to neighborhoods with greater access to preventive care. If this is true, we would expect to see differential trends prior to moving for movers who relocate to high preventive care utilization areas vs. movers who relocate to low preventive care utilization areas. I test for these patterns by examining trends in having an annual PCP visit for movers moving to high and low utilization neighborhoods and find no evidence for these predictions.

The assumption could also be violated if mobility is related to idiosyncratic match effects between enrollees and neighborhoods. For example, if enrollees tend to move to neighborhoods that are particularly conducive to their own preventive care utilization, the estimated gains for movers will overstate gains for a typical enrollee. An implication of this selective mobility is that moves to neighborhoods with expected losses in preventive care utilization would be offset by an improvement in match effects. In the limit, if all moves were driven solely by match components, *all* moves would lead to gains in utilization. I test for signs of selective mobility by examining movers moving in opposite directions between groups of high and low preventive care utilization areas. I find that the utilization changes are approximately symmetrical (i.e. equal in magnitude and opposite in sign), as predicted by the additive model with exogenous mobility.

Finally, the model implies that variation across places causes a level shift in preventive care utilization that does not depend on patient fixed effects α_i or other observables. I find this to be a reasonable assumption, which is strengthened by the immediate and symmetric utilization changes observed with moves across high and low utilization areas. The specification also rules out shocks to utilization that coincide exactly with the timing of the move and that are correlated with utilization in the origin and destination. If this is the case, we might expect a postmove spike in utilization that dissipates over time, and we do not find evidence for this.

1.3.3 Normalization

The model in Equation 1.3 is only identified if the data include movers. If no patients are observed moving across neighborhoods, there would be no way to separate differences in neighborhood fixed effects γ_j from differences in average patient characteristics within each neighborhood. The key quantity for identifying neighborhood fixed effects is the observed change in utilization when patients move.

For the main analysis, I estimate Equation 1.3 separately by patient race. For each sample, the model is a two-way fixed effect model in the style of AKM, with individual and neighborhood fixed effects estimated across panel data. The neighborhood effects in each two-way fixed effects model are only identified within a "connected set" of neighborhoods linked by patient moves. I therefore restrict estimation to enrollees and neighborhoods in the largest connected set for movers of each race.

The analysis produces sets of neighborhood effects $\hat{\gamma}_j^{white}$ and $\hat{\gamma}_j^{black}$. Given that the effect for any given neighborhood is only identified relative to a reference place or set of places, I demean each set of fixed effects, so that the estimates $\tilde{\gamma}_j^{white}$ and $\tilde{\gamma}_j^{black}$ represent the effect of living in neighborhood j relative to the (population-weighted) average neighborhood for enrollees of race r.

Because our goal is to decompose the observed Black-white gap in PCP visits, I do not restrict to a common support of neighborhoods that are in both the Black and white sets of estimated neighborhood effects; if a neighborhood j is estimated only in the Black patient model because we only observe Black patients in the given neighborhood, we would still like to include the contribution of its causal effect to the raw gap.

Still, the sets of $\tilde{\gamma}_j^{white}$ and $\tilde{\gamma}_j^{black}$ are not identified relative to each other. In order to identify the level shift between $\{\tilde{\gamma}_j^{white}\}$ and $\{\tilde{\gamma}_j^{black}\}$ we would need to observe enrollees change race, to examine how utilization changes when patient race changes but neighborhood remains the same.

In the analysis, I normalize $\{\tilde{\gamma}_j^{black}\}$ such that the set of places in the lowest decile of racial segregation have $E[\tilde{\gamma}_j^{white} - \tilde{\gamma}_j^{black}] = 0$. I show empirically that neighborhoods with greater racial segregation exhibit divergence between $\tilde{\gamma}_j^{white}$ and $\tilde{\gamma}_j^{black}$, lending support to this choice of normalization. If this set of places actually has lower causal effects for Black than white residents, my estimate will be a lower bound of the difference in neighborhood effects between Black and white residents. It is unlikely that this set of places would have higher causal effects for Black than white residents.

1.3.4 Decomposition

With the estimated γ coeeficients from Equation 1.3, I decompose the difference in place effects into sorting and match components following the framework of the Oaxaca-Blinder wage decomposition. Letting the notation $\Delta k = E[k|white] - E[k|black]$, we have:

$$\Delta y = \Delta\alpha + \Delta\gamma + \Delta\tau + \Delta x \tag{1.4}$$

$$= (\Delta\alpha + \Delta x + \Delta\tau) + \underbrace{\sum_{j\in J} \gamma_j^{white}(S_j^{white} - S_j^{black})}_{\text{sorting}} + \underbrace{\sum_{j\in J} (\gamma_j^{white} - \gamma_j^{black}) S_j^{black}}_{\text{efficacy}}, \tag{1.5}$$

where S_j^{white} is the share of white beneficiaries living in neighborhood j out of the total number of white beneficiaries observed in the data, and S_j^{black} is the share of Black beneficiaries in neighborhood j out of the total number of Black beneficiaries observed in the data.

The sorting and efficacy components together constitute the portion of the Black-white gap attributable to immediate effects of residential neighborhood. The unexplained portion of the gap can be due to mean Black-white differences in individual preferences and other individual time-varying characteristics.

The first term, the estimated sorting effect, in Equation 1.5 is invariant to the choice of normalization, but the efficacy term depends entirely on the normalization. Because the quantities of interest, the sorting and efficacy terms, are weighted means of the model parameters estimated from OLS, and those estimates are unbiased and normally distributed given our identification assumptions, we can produce an unbiased estimate of the sorting and efficacy terms by inputting the sample analog. On the other hand, estimates based on the parameters themselves, such as the variance of γ_j^{white} or γ_j^{black}, will be biased and would require empirical Bayes or shrinkage estimators.

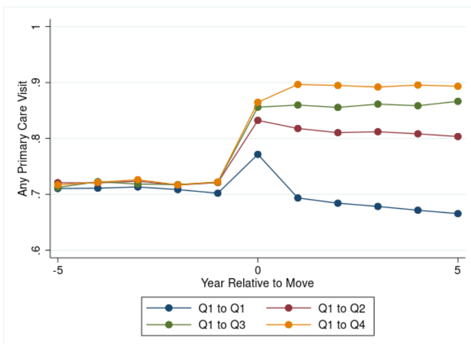
1.4 Results

1.4.1 Descriptive Evidence of Neighborhood Effects

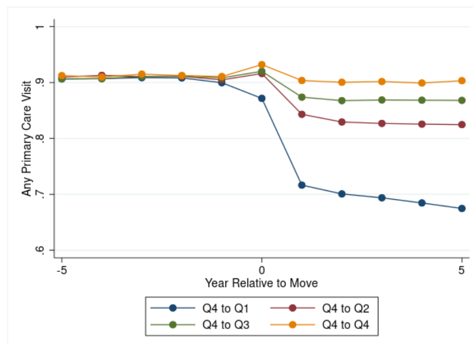
I begin by showing evidence supporting the empirical strategy and identifying assumptions for the two-way fixed effects model described in Equation 1.3. Following Card *et al.* (2013) and Card *et al.* (2016), I plot descriptive evidence on the patterns of having an annual PCP visit for enrollees who move between neighborhoods with higher and lower rates of PCP visits among non-moving residents. In doing so, I document several facts that are consistent with the additivity and exogenous mobility assumptions in the model setup.

Using data on movers within the five years before and after their move year, I categorize each mover into quartiles based on their origin and destination neighborhood PCP visit rates, as measured by annual PCP visit rates among stayers. This generates sixteen groups of origin and destination quartile pairs. I show trends in having an annual PCP visit, before and after the move, for movers with origin neighborhoods in the lowest and highest quartile of PCP visit rates among stayers in Figure 1.5. Results for the white mover sample are

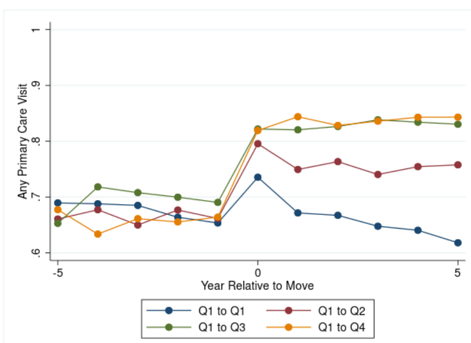
Figure 1.5: *Changes in Primary Care Visits Around Moves*



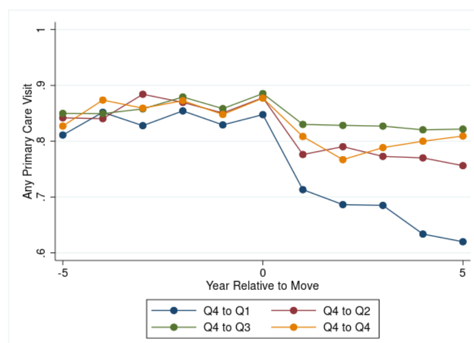
(a) *Low Origin PCSA, White Movers*



(b) *High Origin PCSA, White Movers*



(c) *Low Origin PCSA, Black Movers*



(d) *High Origin PCSA, Black Movers*

Notes: Figures show the fraction of Medicare enrollees with any primary care visit in the five years before and after a move, separated by enrollee race and type of move. The sample restricts to Medicare enrollees who moved exactly once during 1999 to 2017. PCSAs are separated into quartiles by their observed primary care visit rates among Medicare stayers.

shown in the top panel, and results for the Black mover sample are shown in the bottom panel.

First, I observe that enrollees who move between neighborhoods with higher and lower rates of PCP visits among stayers experience systematic gains and losses in likelihood of having an annual PCP visit. This suggests that there are significant neighborhood-specific effects on primary care take-up, for both Black and white enrollees. Second, there is no evidence that movers to destination neighborhoods with higher primary care take-up experience differential trends in primary care prior to their move. Third, changes in primary care take-up appear only in the year of and the year following the move, with no evidence of differential trends in primary care following the move which might be expected if an evolving health shock prompted the move.

Fourth, the gains and losses from moving between neighborhoods with higher and lower rates of PCP visits are approximately symmetric, suggesting that neighborhood effects are additively separable from other drivers of primary care take-up and are not driven by selective mobility and idiosyncratic match effects. Finally, Black movers seem to gain less than white movers from moving to high PCP rate neighborhoods. This is consistent with a model with heterogeneous neighborhood effects by race, driven by differences in within-neighborhood "efficacy" of primary care.

1.4.2 Estimation of Patient-Neighborhood Model

Building on the natural experiment of moving from a low PCP visit to a high PCP visit neighborhood, I estimate a full two-way fixed effect model with individual and neighborhood fixed effects, separately by patient race. I define neighborhoods as Primary Care Service Areas (PCSAs), which are aggregations of zip codes such that the majority of patients living in an area use primary care services from within the area.

Restricting to patient-year observations within the five years before and after the move year, and omitting the move year, I measure the extent to which patients' PCP visit probability changes upon move from one PCSA to another, for the largest connected set of PCSAs in the movers samples for each race.

Table 1.2 shows that the analysis sample includes 6,633 PCSAs for the white movers estimation and 2,337 PCSAs for the Black movers estimation. Given the large number of fixed effects in the model, the parameters α_i and γ_j^r are estimated with considerable noise, especially where the sample size as measured by the number of observations in the data in neighborhood j and race r is small. I follow Kline *et al.* (2020) to implement a leave-out correction for the estimate of the variance in neighborhood effects for the Black and white patient samples.

Indeed, the plug-in estimates of the standard deviation of the distribution of α_i and γ_j^r are larger for the Black patient sample, where the sample size is smaller and the parameter estimates are more noisy. However, after implementing the leave-out correction, the distributions of person and PCSA effects appear similar for the Black and white samples.

The estimated standard deviation in neighborhood effects is 8-10pp, which measures the expected change in annual PCP visit probability from moving an individual from one neighborhood to another neighborhood with 1 SD higher causal effect on PCP visit rates. This magnitude is greater than the raw Black-white disparity in probability of having an annual PCP visit, but is smaller than the raw geographic variation in PCP visit rates across neighborhoods. Consistent with prior work on geographic variation in health care utilization, I find that the observed geographic variation reflects both place-specific effects as well as some selection on individual characteristics of residents.

Figure 1.6 shows the estimated $\hat{\gamma}_j^r$, binned by deciles of $\hat{\gamma}_j^{white}$. Although the sets of place effects are positively correlated, the relationship is not one-to-one.

1.4.3 Neighborhood Effects and the Black-white Gap

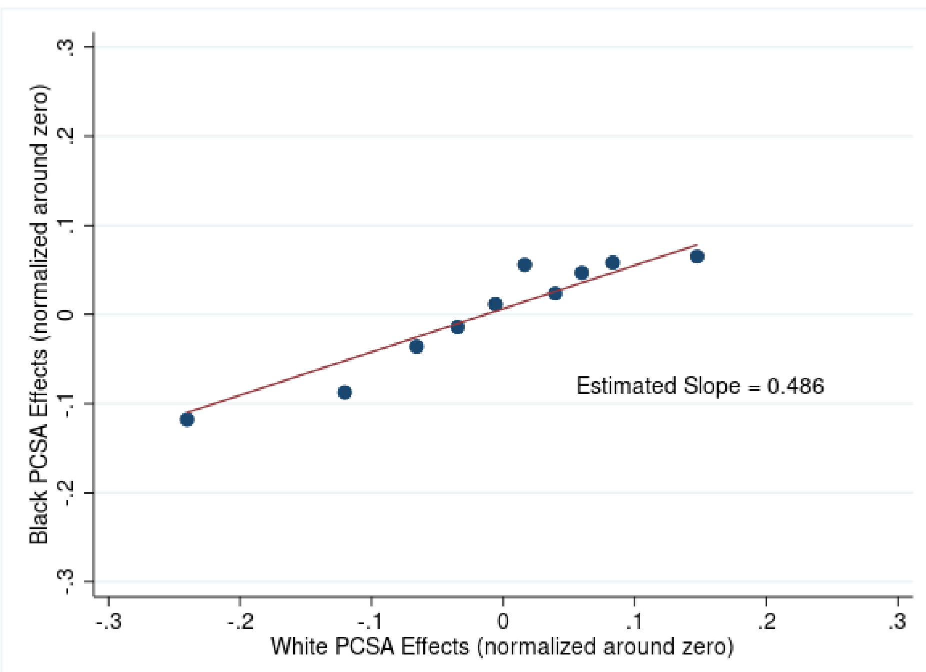
I use the estimated neighborhood effects $\hat{\gamma}_j^r$ to construct a decomposition of the Black-white gap into place-based sorting and efficacy components, as described in Section 1.3.4. First, in order to normalize the sets of $\hat{\gamma}_j^{black}$ and $\hat{\gamma}_j^{white}$ relative to each other, to quantify their difference, I construct a measure of residential racial segregation at the neighborhood level.

Table 1.2: *Summary of Estimated TWFE Models by Race*

	White	Black	Diff
A. Connected set:			
Mean	0.823	0.753	0.070
Number of person-year observations	2,058,164	96,867	
Number of person effects	266,114	12,615	
Number of PCSA effects	6,633	2,337	
B. Summary of parameter estimates:			
Plug-in: SD person effects	0.256	0.337	
Plug-in: SD place effects	0.109	0.181	
Bias corrected: SD person effects	0.157	0.211	
Bias corrected: SD PCSA effects	0.096	0.083	

Notes: Panel A shows summary statistics for the connected sets of Medicare enrollees who moved exactly once within 1999 to 2017, used to estimate the two-way fixed effect models. Panel B shows a summary of the estimated fixed effects, using leave-out estimation as in Kline et al. (2020) to correct for bias in the variance of the parameter sets.

Figure 1.6: *PCSA Effects by Race for Joint Connected Set*



Notes: Figure shows mean estimated PCSA effects binned into 10 deciles by estimated White enrollee PCSA effects on having a primary care visit. PCSA effects estimated on connected set of Medicare enrollees who move once during 1999-2017 for each race. Figure restricts to PCSAs in joint connected set.

I use a Theil index with higher values indicating greater racial segregation, constructed by considering deviations in the racial make-up of census tracts within a PCSA relative to the racial make-up of the PCSA as a whole.

Figure 1.7 shows means of the estimated neighborhood effects $\hat{\gamma}_j^r$, separately by race and binned by deciles of local racial segregation. The figure shows that, despite PCSAs being defined as units of primary care services, PCSAs with greater local racial segregation have particularly worse causal effects on average for Black patients relative to white patients. I normalize the Black PCSA effects relative to the white PCSA effects such that the best fit lines intersect for the decile of PCSAs with the least racial segregation.

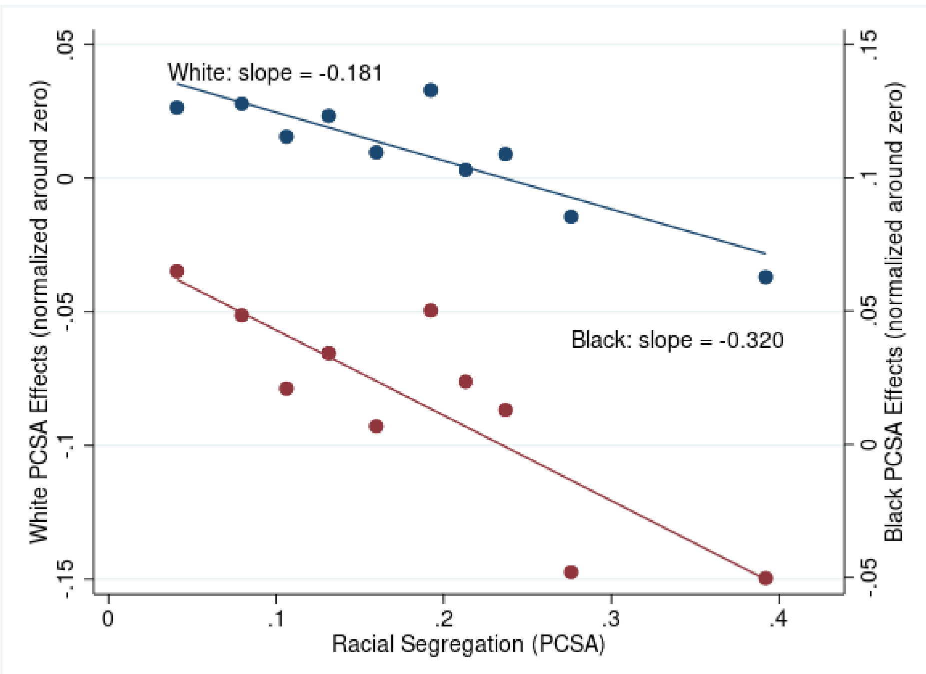
Following this normalization, Table 1.3 presents the estimated sorting and efficacy components of Equation 1.5. I report the estimated contribution of each component in Columns (4) and (5), with the contribution as a percent of the Black-white gap reported below each estimate in parentheses. For the full sample of movers in Panel A, the sorting component contributes 28% to the Black-white gap in PCP visits, and the efficacy component contributes 16% to the Black-white gap in PCP visits. The total contribution of place-based effects is 44%.

The decomposition results are reflected similarly by enrollees across all age groups. Panel B of Table 1.3 shows the Black-white gap and decomposition of place effects by 5-year age bin. Across age bins, the Black-white gap in having an annual PCP visit remains consistent around 8-9pp, with meaningful contributions by both sorting across neighborhoods and efficacy within neighborhoods to generate a total contribution of place effects of 40-50%. The results indicate that, across the sample, Black enrollees consistently live in neighborhoods with lower causal effects on primary care, while also being left behind within neighborhoods.

1.4.4 Robustness of Role of Neighborhood Effects

I conduct two alternative two-way fixed effects specifications, varying the geographical unit defined as a neighborhood. In particular, I estimate the causal effects attributed to counties and hospital referral regions. Table 1.4 shows that, regardless of the definition of

Figure 1.7: *PCSA Effects and Local Racial Segregation*



Notes: Figure shows mean estimated PCSA effects by race binned into 10 deciles of PCSA-level residential racial segregation. Racial segregation is computed using a Theil index measuring the extent to which the racial distribution among census tracts deviates from the overall racial distribution of the PCSA.

Table 1.3: *Contribution of Neighborhood Effects to Black-white Gap in PCP Visits*

	Observed PCP Visit Rates			Dcomposition of Contribution of Place Component		Total Contribution
	White Mean (1)	Black Mean (2)	Black-white Gap (3)	Sorting (4)	Efficacy (5)	of Place Components (6)
A. All	0.836	0.751	0.085	0.024	0.014	0.038
				(28.2)	(16.0)	(44.7)
B. By age group:						
65-69	0.827	0.745	0.083	0.029	0.012	0.040
				(34.8)	(14.1)	(48.9)
70-74	0.851	0.776	0.076	0.025	0.011	0.036
				(33.0)	(14.3)	(47.3)
75-79	0.852	0.764	0.088	0.025	0.014	0.039
				(28.0)	(16.2)	(44.3)
80-84	0.843	0.749	0.094	0.023	0.015	0.038
				(24.0)	(16.4)	(40.4)
85+	0.802	0.709	0.094	0.019	0.018	0.037
				(20.3)	(18.8)	(39.1)

Notes: Table shows results from the decomposition of the Black-white annual primary care provider (PCP) visit gap into components attributable to neighborhood effects. Columns 1-3 show observed annual PCP visit rates for the movers sample by beneficiary race. Columns 4-6 show the estimated percentage point difference (White - Black) attributable to neighborhood effects, overall (Column 6) and through differences in sorting across neighborhoods (Column 4) vs. efficacy within neighborhoods (Column 5). Decomposition is conducted using parameter estimates from TWFE models across PCSAs by enrollee race. Percentage point estimates are reported, with percent of the Black-white gap accounted for reported in parentheses. Panel A shows results for the full sample of enrollees within PCSAs that are in both the connected sets of Black enrollee movers and white enrollee movers. Panel B shows decomposition results by enrollee age group in 5-year age bins.

"neighborhood", place effects account for about 40-45% of the Black-white gap in having an annual PCP visit. As the defined neighborhoods grow larger, from PCSAs to counties and from counties to HRRs, the fraction of the difference in place effects attributable to sorting across neighborhoods shrinks, while the fraction attributable to efficacy within neighborhoods grows.

Defining neighborhoods at the HRR level, we find that sorting across HRRs does not contribute at all to the Black-white gap in having an annual PCP visit: Black enrollees do not on average live in regions of the country with lower causal effects on primary care than white enrollees. However, the total contribution of place effects in the Black-white gap remains at 43%, as Black enrollees are more substantially left behind in HRRs with greater residential racial segregation. These results indicate that regardless of the definition of "neighborhood" residential segregation acts as a meaningful proxy for areas where Black patients are particularly left behind. Further, differences in place effects contributing to the Black-white gap are entirely due to micro-level segregation within regions as opposed to being driven by differences in region of residence at the HRR or broader level.

1.5 Conclusion

Focusing on over-65 Medicare beneficiaries, I show that 40-50% of the Black-white disparity in the presence of an annual primary care visit is due to immediate place- and place-by-race specific factors. The remaining 50-60% of the disparity is due to individual characteristics of enrollees that are carried with them when they move and may differ on average by race. Further, the 44% of the disparity that is generated by place effects for the full sample can be decomposed into 28% due to differences in sorting across primary care neighborhoods and 16% due to differences in place-by-race effects within neighborhoods.

The analysis shows that external place effects play a substantial role in the persistent Black-white gap for an important measure of health care quality. In particular, I conceptualize the roles of across-place sorting and within-place efficacy, an important distinction from prior studies that consider the role of sorting alone. The results differ from the conclusions

Table 1.4: *Sensitivity of Decomposition by Neighborhood Definition*

	Observed PCP Visit Rates			Dcomposition of Contribution of Place Component		Total Contribution
	White Mean (1)	Black Mean (2)	Black-white Gap (3)	Sorting (4)	Efficacy (5)	of Place Components (6)
A. PCSA:						
	0.836	0.751	0.085	0.024	0.014	0.038
				(28.2)	(16.5)	(44.7)
B. County:						
	0.836	0.751	0.085	0.012	0.022	0.034
				(14.1)	(25.9)	(40.0)
C. HRR:						
	0.836	0.751	0.085	-0.002	0.038	0.036
				-(2.4)	(44.7)	(42.4)

Notes: Table shows results from the decomposition of the Black-white annual primary care provider (PCP) visit gap into components attributable to neighborhood effects. Columns 1-3 show observed annual PCP visit rates for the movers sample by beneficiary race. Columns 4-6 show the estimated percentage point difference (White - Black) attributable to neighborhood effects, overall (Column 6) and through differences in sorting across neighborhoods (Column 4) vs. efficacy within neighborhoods (Column 5). Percentage point estimates are reported, with percent of the Black-white gap accounted for reported in parentheses. Panels A-C show results varying the definition of neighborhood, with neighborhoods defined as Primary Care Service Areas in Panel A, counties in Panel B, and Hospital Referral Regions in Panel C. Decomposition is conducted using parameter estimates from TWFE models across PCSAs, counties, and HRRs by enrollee race.

that would be drawn without allowing for heterogeneous place effects by race.

These findings add to the conclusions of the existing literature on racial disparities in health care. I find broad evidence that, within-place, Black patients face barriers to receiving health care relative to white patients. I also build on work regarding residential segregation, finding consistent evidence for sorting across neighborhoods driving some share of the Black-white gap. By considering different geographies for place-based effects, I test both macro- and micro-segregation channels and find that micro-segregation is the dominating driver of racial disparities in this setting.

The results suggest that future research and policy should focus on understanding causes and correlates of local-level residential racial segregation, to improve access to equal health care environments. More broadly, I show that, among a 65 and older insured population, Black-white disparities in basic health care continue to persist, and will require a multi-faceted approach to address. Individual differences, neighborhood differences, and within-neighborhood differences all play substantial roles in shaping the existing Black-white gap in primary care utilization.

Chapter 2

Heterogeneous Hospital Quality by Race: Evidence from Emergency Events

Abstract

Measures of hospital quality, such as risk-adjusted mortality and patient satisfaction, may mask heterogeneous treatment effects across patients, with implications for population-level health disparities and hospital choice. In this paper, I test whether there are differences in the ordered quality ranking of local hospitals for Black and white emergency patients. To control for underlying factors that affect hospital choice, I exploit ambulance company preferences as an instrument for hospital characteristics and estimate the impact of being treated at a hospital with a higher share of patients who are Black. Using mortality following hospitalization as the primary outcome, I find substantial race-specific effects, with Black patients experiencing better outcomes at hospitals with greater Black patient shares in the medium to long run. The results are consistent with a hospital choice model that exhibits institutional comparative advantage and positive Roy selection by patient race, pointing to the limitations of broad-based hospital quality measures.

2.1 Introduction

There is strong interest in measuring the quality of services rendered in health care, in order to improve efficiency and reduce soaring health care costs. In line with these goals, measures of hospital quality including risk-adjusted 30-day mortality and hospital "report cards" have become widely used in policymaking to reward high-performing hospitals with higher reimbursement rates and to influence both providers and patients.

Recent literature has shown that measures of hospital quality largely correspond to quasi-experimental estimates that address selection bias, but has also documented the presence of consequential selection-on-gains within the patient-hospital choices observed (Doyle *et al.*, 2019; Hull, 2020). The presence of selection-on-gains implies that hospitals specialize in patient types while patients are somewhat aware of this institutional comparative advantage. Given that patients are admitted to hospitals that are idiosyncratically better for them, existing hospital quality measures have limited ability to improve patient outcomes through informing hospital choice.

In this paper, I test for the presence of selection-on-gains by patient race, in particular for Black Medicare patients. It has long been documented that Black patients are more likely to receive care at safety-net hospitals and hospitals with lower quality scores than white patients, even when living in the same zip code and despite actually living closer to high-quality hospitals than white patients on average (Dimick *et al.*, 2013; Hanchate *et al.*, 2019; Chandra *et al.*, 2020). These observations have been used to encourage health care providers to steer Black patients towards hospitals measured as higher-quality.

In contrast, regarding existing hospital choice patterns as suggestive evidence of potential selection on heterogeneous gains, I develop an instrumental variables (IV) framework to test whether Black patients have better outcomes at hospitals that are seen to treat more Black patients. Specifically, I construct estimates of the impact of an emergency patient being admitted to a hospital with a higher Black patient share, for Black and white patients separately. In order to produce causal estimates, I leverage the quasi-random assignment of patients to ambulance companies with "preferences" for certain hospitals to instrument for hospital characteristics. I implement this empirical strategy on an analysis sample of Medicare patients who are admitted to a hospital for a “non-discretionary,” emergency condition, and who are transported by ambulance.

The ambulance-instrument approach relies on the assumptions of exogeneity of ambulance company assignment, relevance of ambulance assignment for determining hospital characteristics, and an exclusion restriction that ambulance assignment affects patient outcomes only through hospital choice. Prior work, beginning with Doyle *et al.* (2015) and subsequently adapted by others, has shown plausibility of quasi-random assignment of ambulance companies when controlling for patient zip code of residence and type of pickup origin (e.g. at home, in a nursing home, or at the scene of an accident or illness). I follow prior work to construct ambulance IVs corresponding to hospital Black patient shares and total patient volume, and find plausible balance along observed patient characteristics as well as strong first-stage coefficients.

The two-stage least squares (2SLS) estimates show that Black patients assigned to hospitals with higher Black patient shares have lower mortality in the 180 days following a hospital admission. The magnitude of the effect is large, implying that being admitted to a hospital with a 1 SD higher Black patient share among local hospitals reduces 180-day patient mortality on average by 0.77-0.95 percentage points, on a mean of 30.8 percent. The effect is of a similar magnitude to being admitted to a hospital with a 1 SD higher patient volume.

On the other hand, there is no effect of being admitted to a hospital with a higher Black patient share for white patients. Using patterns for white patients as a control group for Black patient outcomes, I construct an IV differences-in-differences specification, which shows the same pattern of reduced mortality for Black patients, with slightly larger coefficients and stronger statistical significance. Even in reduced form specifications, the race-specific patterns are striking and are distinct from conditional OLS estimates. The contrast between conditional OLS and reduced form coefficients for Black patients implies selection bias of sicker Black patients to hospitals with a higher share of Black patients and demonstrates the importance of using quasi-random variation to understand hospital effects.

The results from this paper contribute to a literature positing the importance of institutional comparative advantage and productivity spillovers in determining treatment choices, patient outcomes, and overall efficiency in health care (Chandra and Staiger, 2007, 2020). For example, hospitals may invest in technologies, programs, and health care workers who specialize in specific diagnoses and are more valuable for certain patient groups, generating complicated predictions about optimal patient sorting and levels of care. Specifically in the context of this paper, Black patients could have better average outcomes at hospitals that treat more Black patients if there is a higher probability of seeing a Black physician or being referred to clinics, specialists, or other follow-up care providers that facilitate trusted and open patient-physician communication for Black patients (Alsan *et al.*, 2019; Hill *et al.*, 2020).

The results are consistent with a Roy model of institutional choice, containing self-selection by rational actors making optimizing decisions about what markets to participate in - in this case which hospital to go to. In the sense of Roy (1951), we would expect the individuals most likely to select an institution to see systematically higher gains from doing so. Productivity spillovers among Black patients would further reinforce the presence of comparative advantage and positive Roy selection across hospitals. These insights caution against the assumption that particular hospital choice patterns reflect poor decision-making and highlight the limits of using broad-based hospital quality measures to improve patient outcomes.

2.2 Data

To assess heterogeneous hospital quality by race, I use data from a 20 percent sample of Medicare beneficiaries in the years 2003 to 2014. Given the scope of public health insurance, the Medicare data are nationally representative among individuals age 65 and older and include patients treated at the vast majority of hospitals and providers in the United States. Further, beneficiaries are observed from year to year and can be linked to long-term outcomes including mortality.

2.2.1 Sample Construction

Following Doyle *et al.* (2015), I identify patients who are admitted to an acute care hospital for a non-discretionary condition after being transported by ambulance to the emergency department.

I define "non-discretionary" conditions as the set of 29 3-digit ICD-9 principal diagnosis codes with weekend admission rates as close or closer to 2/7ths as hip fracture, reflecting a lack of discretion in the timing of the hospital admission[1]. These non-discretionary conditions are likely to be serious and require immediate care, serving as a sample with which the ambulance instrument is well matched. In particular, the empirical strategy relies on random assignment of patients to ambulance companies, which is plausible in settings where the timing of and need for transport is sudden and non-negotiable. I identify the set of non-discretionary inpatient hospital admissions using the Medicare Provider Analysis and Review (MEDPAR) files.

I then use the Medicare carrier and outpatient claims files to identify ambulance transports. The carrier file contains the large majority of ambulance charges, including claims submitted by organizational ambulance providers, and the outpatient file contains ambulance claims that are affiliated with a hospital or other facility charge. Ambulance claims are identified using the reported place of service code and HCPCS modifier codes for ambulance services[2]. I link the ambulance transport data to inpatient admissions using the patient identifier and the date of hospital admission.

Bringing these sets together, I construct the analysis sample to include hospital admission events that are linked to an ambulance transport, have an emergency department charge,

[1]The full set of principal diagnosis categories included as non-discretionary conditions is listed in Appendix Table A1 of Doyle *et al.* (2015), along with their weekend admission rates.

[2]Ambulance claims have a place of service code of 41 for land ambulances and a HCPCS modifier code of RH, SH, IH, EH, NH, JH, or PH, which indicate that the patient is transported to a hospital from a residence, the scene of a accident, a transfer site, a nursing home, a skilled nursing facility, a non-hospital-based dialysis facility, or a physician's office, respectively.

and have one of the 29 non-discretionary conditions as the primary ICD-9 code reported for the admission. If a beneficiary appears multiple times in the data set, I consider only the first event. In addition, I restrict the sample based on information in the Medicare enrollment file, to beneficiaries that were enrolled in Medicare Parts A and B, with no HMO enrollment, for 12 months before and after the hospitalization, or until death. This ensures that the sample is comprised of patients whose hospital and physician charges are observed in the fee-for-service data.

For the resulting sample of ambulance admission events I link to patient demographic and other characteristics in the Medicare enrollment file, including age, race, sex, and zip code of residence[3]. I construct indicators for beneficiary race (Black, white, other), 5-year age groups, and sex, to be used as control variables in the regression analysis. I also generate indicators for 3-digit ICD-9 principal diagnosis codes, the year of the hospital admission event, and 17 comorbidities constructed by mapping each patient's claims in the year prior to (but not including) the admission event to hierarchical condition codes (HCC).

I construct a set of ambulance-related variables following prior work. These include miles traveled with the patient, whether the ambulance used emergency lights and sirens, whether the ambulance has advanced life support capabilities, and the ambulance payment amount. I use linked vital statistics data to construct the primary outcome variables of 180-day and 360-day mortality from the date of hospital admission.

Finally, before conducting analysis, I restrict to zip codes with at least one Black and white patient event, hospitals with at least 30 admissions in the sample, and ambulance companies with at least 20 transports. For the main analysis, I focus only on Black and white patient samples, dropping records of other races. In total, this yields 745,000 ambulance admission events in the final sample.

[3] The race variable is recorded from self-reported Social Security Administration data. Zip code of residence specifies the address on file where beneficiaries receive cash benefits, including Social Security.

2.2.2 Measures of Patient Volume by Race and Hospital

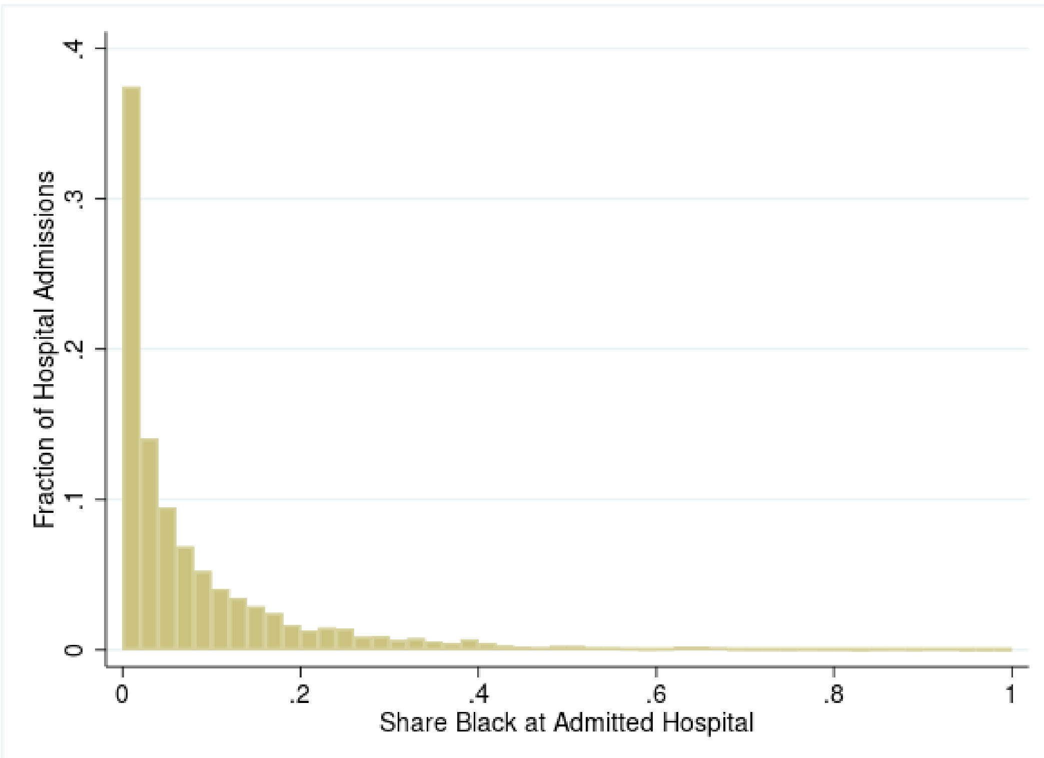
I construct hospital-level measures of patient volume and racial demographics using the 20 percent sample of hospital admissions for non-discretionary conditions in the MEDPAR files, without restricting to ambulance transports. Using these data from 2000 to 2014, I compute counts of Black patient admissions, white patient admissions, and patient admissions of other races for each acute care hospital and year. In particular, I measure two key hospital characteristics for the analysis: total patient volume and the share of patients who identify as Black.

For every year in the analysis period of 2003 to 2014, I assign to each hospital the hospital characteristic as measured by its 3-year lagged mean. For example, for the year 2007, I measure the hospital's patient experience in that moment as the mean patient volume and racial demographics of the prior three years, 2004 to 2006. This lagged construction of hospital characteristics ensures that I am not including an analysis patient himself in his corresponding measure of hospital experience, nor any patients who are admitted to the hospital in future periods. I compute a 3-year mean to smooth out idiosyncratic year-to-year variation in patient admissions.

Figure 2.1 shows the distribution of the hospital share Black variable, averaged by hospital across years 2003-2014 and weighted by patient volume. The hospital measures of patient volume and racial composition are highly correlated across years. In the inpatient data used to construct the hospital measures, the median patient is admitted to a hospital with a Black patient share of 3.7 percent.

Given the skewed distribution, for the main analysis I take the log of both hospital share Black and total volume; regression coefficients on the hospital characteristics are then interpreted as the effect of percent increases from baseline. Because the hospital characteristics are constructed solely from Medicare patient samples, they may not reflect overall racial demographics if we were to consider younger patients as well. However, the analysis sample of Medicare ambulance transports is highly comparable to the broader Medicare patient pool, so these hospital characteristics serve as relevant measures for

Figure 2.1: *Distribution of Hospital-level Black Patient Share*



Notes: Figure shows variation in the share of Black patients at admitting hospitals for Medicare beneficiaries admitted to acute care hospitals with non-discretionary conditions between 2003 and 2014. Histogram includes 50 bins.

potential productivity spillovers.

2.3 Empirical Strategy

The empirical approach compares long-run mortality outcomes of Medicare emergency patients quasi-randomly assigned to hospitals with higher and lower Black patient shares as a fraction of total patient volume. Following existing work, I use information about the ambulance companies that are quasi-randomly assigned to transport patients in the sample to construct instrumental variables (IVs) capturing exogenous variation in hospital type.

Specifically in this paper, I measure the extent with which particular ambulance companies are likely to take patients to hospitals with higher or lower Black patient shares. I use the ambulance company propensities to instrument for the hospital characteristic, allowing me to construct a two-stage least-squares (2SLS) estimate of the causal effect of being treated at a hospital with higher or lower Black patient share.

2.3.1 Ambulance Referral Patterns in Prior Work

Prior work has established the feasibility of and a methodology for using ambulance company assignment to instrument for hospital choice. As discussed in detail in e.g. Doyle *et al.* (2015), Doyle *et al.* (2019), and Hull (2020), ambulance company assignment is effectively random among companies that serve a particular location, while having an influence on hospital assignment. Common mechanisms for ambulance assignment include rotational assignment, where the 911 dispatch mechanism rotates between ambulance services within a community, and direct competition, where the call for service is broadcast to multiple companies and whichever arrives first gets the business. Both mechanisms generate randomness in ambulance company assignment based on the idiosyncratic rotation point or locations of ambulance providers at the time of an emergency dispatch. Further, ambulance companies exhibit "preferences" for specific hospitals, due to their ownership, affiliation, or base location of operations.

These insights have been operationalized using Medicare fee-for-service claims data, with tests for quasi-random assignment and relevance of ambulance company identifiers. Prior work has found plausible balance along observed patient characteristics across ambulance company assignment, conditioning on the patient zip code and type of origin (e.g. at home, in a nursing home, or at the scene of an accident or illness). In other words, these analyses show that among emergency patients picked up in the same zip code and origin type, there are no systematic differences in age, gender, or existing comorbidities between those picked up by different ambulance companies.

For the IV estimation and set-up, prior work has used ambulance company identifiers to instrument for hospital choice and answer questions about hospital characteristics in two broad ways. One approach has been to consider a standard linear approximation of the relationship between a particular hospital characteristic and patient outcomes. In this framework, the variation in ambulance companies is projected onto a one-dimensional measure of how likely the patient is to be taken to a hospital with a particular characteristic. The 2SLS estimation utilizes as many instruments as there are hospital characteristics in the main regression of interest, generating relatively precise and interpretable IV estimates.

Studies using this approach have found that higher-spending hospitals (Doyle *et al.*, 2015), hospitals with better quality measures as defined by CMS Hospital Compare (Doyle *et al.*, 2019), hospitals operated by the US Department of Veterans Affairs (Chan *et al.*, 2021), and higher-priced hospitals (Cooper *et al.*, 2022) reduce patient mortality on average. One limitation of this approach is that other hospital characteristics correlated with the characteristic of interest could be driving the observed causal effect, and remain undetected.

Alternatively, Hull (2020) employs the full set of ambulance company identifiers as instruments to estimate quasi-experimental 30-day mortality outcomes for each hospital. Because of the large number of endogenous variables (hospitals) and instruments (ambulance companies) involved in this approach, Hull implements a semi-parametric shrinkage methodology to estimate quality posteriors for each hospital with meaningful precision. With this additional structure, Hull is able to estimate correlations between a range of hospital characteristics and quasi-experimental hospital performance, and shows that higher-spending, higher-volume, and privately owned hospitals reduce mortality on average. These findings are consistent with other literature using the ambulance-instrument approach.

In addition, Hull finds that most markets exhibit positive Roy selection on heterogeneous gains. In other words, patients tend to be admitted to more appropriate hospitals in markets with hospital comparative advantage. Given this positive Roy selection, Hull finds that the policy implications of using hospital quality rankings to inform hospital choice are limited.

2.3.2 Estimating the Effect of Hospital Black Patient Share

Building on existing literature, I test for Roy selection along patient race for Black and white Medicare patients. The presence of positive selection-on-gains could explain differences in hospital choice patterns by race and would have implications for policies to reduce racial disparities in hospital outcomes.

The empirical strategy uses 2SLS regression including two endogenous variables and two instrumental variables, to estimate the effect of being admitted to a hospital with a higher or lower Black patient share, while controlling for total patient volume. The coefficient of interest is the relationship between a hospital's Black patient share and patient outcomes, but total patient volume is included as a second endogenous variable to control for the fact that higher volume hospitals have been shown to have better outcomes (Hull, 2020). To test for hospital comparative advantage and productivity spillovers by race, I estimate separate regressions for Black and white patient samples.

There are several possible sets of results, which would have different implications. Being admitted at a hospital that sees more Black patients could increase (decrease) 180-day and 360-day mortality for both Black and white patient samples, indicating that these hospitals are worse (better) for both patient groups. Similarity of the magnitudes of these estimates would suggest that across race on average, hospital quality is consistent and not heterogeneous. On the other hand, the coefficients could differ by patient race, implying the presence of race-specific hospital quality. If there are positive selection-on-gains by race, we would expect to see reduced mortality for Black patients assigned to hospitals with more Black patients, and an increase or no effect on mortality for white patients assigned to hospitals with more Black patients.

Using the analysis sample of Medicare emergency patients transported by ambulance, I create a measure of ambulance propensities to take patients to certain types of hospitals. The hospital characteristics I instrument for are the share of patients who identify as Black, $HospB_{ht}$, and the total patient volume, $HospV_{ht}$, for hospital h in year t, constructed as 3-year lagged means and log-transformed as described in Section 2.2.2. Then, for a given patient i assigned to ambulance $a(i)$, I instrument for hospital characteristics $HospB_{h(i)t(i)}$ and $HospV_{h(i)t(i)}$ using the leave-out mean of the realized hospital characteristics for other patients j, whose total is denoted $N_{a(i)}$, picked up by the same ambulance company:

$$Z_i^B = \frac{1}{N_{a(i)} - 1} \sum_{j \neq i}^{N_{a(i)}-1} HospB_{h(j)t(j)} \tag{2.1}$$

$$Z_i^V = \frac{1}{N_{a(i)} - 1} \sum_{j \neq i}^{N_{a(i)}-1} HospV_{h(j)t(j)}. \tag{2.2}$$

The two instruments Z_i^B and Z_i^V represent ambulance company tendencies to take patients to hospitals with the corresponding characteristics.

I use these instruments to estimate the first-stage relationship between the characteristics of the hospitals where patients are admitted, $HospB_i$ and $HospV_i$, and the instruments for these characteristics, Z_i^B and Z_i^V. This yields two first-stage regressions for the two endogenous hospital characteristics, estimated for patient i conditional on principal diagnosis group $d(i)$, year of event $y(i)$, and zip code of residence by origin type of pickup $z(i)$:

$$HospB_i = \check{\alpha}_0 + \check{\alpha}_1 Z_i^B + \check{\alpha}_2 Z_i^V + \check{\alpha}_3 X_i + \check{\alpha}_4 A_i + \check{\gamma}_{d(i)} + \check{\theta}_{z(i)} + \check{\lambda}_{y(i)} + \check{v}_i \tag{2.3}$$

$$HospV_i = \tilde{\alpha}_0 + \tilde{\alpha}_1 Z_i^B + \tilde{\alpha}_2 Z_i^V + \tilde{\alpha}_3 X_i + \tilde{\alpha}_4 A_i + \tilde{\gamma}_{d(i)} + \tilde{\theta}_{z(i)} + \tilde{\lambda}_{y(i)} + \tilde{v}_i. \tag{2.4}$$

Following prior work, X_i is a vector of patient controls including 5-year age bins, sex, and indicators for seventeen common comorbidities, and A_i is a vector of ambulance transport-related controls including distance traveled in miles, utilization of advanced life support (e.g. paramedic) capabilities, use of emergency traffic signals (e.g. sirens and lights), and payment amount. The ambulance controls are particularly important to consider, as the exclusion restriction would be violated if ambulance company assignment affects patient outcomes directly, not just through its effect on hospital choice. I address this concern by reporting specifications with and without the ambulance controls. I cluster standard errors at the hospital service area (HSA) level, given that local market areas may have their own assignment rules.

Finally, I estimate the 2SLS relationships between hospital characteristics and mortality, leveraging quasi-random variation in ambulance assignment. With the predicted characteristics $HospB_i$ and $HospV_i$ from the first stage regressions, I estimate the second stage regression including the same set of control variables:

$$Y_i = \beta_0 + \beta_1 \widehat{HospB}_i + \beta_2 \widehat{HospV}_i + \beta_3 X_i + \beta_4 A_i + \gamma_{d(i)} + \theta_{z(i)} + \lambda_{y(i)} + \nu_i. \tag{2.5}$$

Outcomes Y_i include patient mortality 180 days and 360 days following the date of hospital admission. I consider other mortality intervals in additional analysis. I measure patient mortality as the main outcome because it serves as a broad indicator of patient outcomes, is reliably observed in the data, and can be clearly interpreted as undesirable.

The key coefficient of interest in Equation 2.5 is β_1. The interpretation is that going to a hospital with a 1% increase in Black patient share impacts mortality by approximately $\hat{\beta}_1$ percentage points. The identifying assumptions for the 2SLS estimate are that ambulance company assignment is random conditional on the included set of controls, including patient zip code by origin type, and affects mortality only through hospital choice. I assess the plausibility of the random assignment assumption across observable dimensions in Section 2.4.1 and the strength of the first stage in Section 2.4.2. To address concerns about the exclusion restriction, I consider the reduced form relationship between ambulance propensity and patient mortality, as well as specifications with and without ambulance controls.

2.3.3 IV Differences-in-differences Specification

In addition, I construct a differences-in-differences estimate of the effect of hospitals' Black patient share for Black patients, using white patient outcomes as a control group. To do this, I interact all regressors with an indicator for the patient being Black and estimate the 2SLS model across the entire sample of Black and white patients combined. This IV differences-in-differences specification includes four instruments (and four first stage regressions, one corresponding to each endogenous variable). The second stage regression is:

$$\begin{aligned} Y_i = & b_0 + b_1 \widehat{HospB}_i + b_2 \widehat{HospV}_i + b_3 K_i + \theta_{z(i)} \\ & + c_1 Hosp\widehat{B_i \times B}lack_i + c_2 Hosp\widehat{V_i \times B}lack_i + c_3 K_i \times Black_i + \epsilon_i. \end{aligned} \tag{2.6}$$

Controls K_i include the set of controls in the prior specifications – $X_i, A_i, \gamma_{d(i)}$ and $\lambda_{y(i)}$ – which are all interacted with $Black_i$, an indicator for the patient being Black.

In addition to the IV identifying assumptions in Section 2.3.2, the differences-in-differences specification requires the additional assumption that the observed outcomes for white patients represent the potential outcomes of Black patients affected by the ambulance company propensity if they were white. This assumption is supported by similar patterns of relationships between patient covariates and the instruments within the Black and white patient samples, in tests of instrument exogeneity in Section 2.4.1.

2.4 Results

2.4.1 Balance

First, I test the plausibility of the exogeneity assumption from the 2SLS framework set up in Section 2.3.2. To do this, I regress a set of observed patient demographic and health characteristics on the constructed ambulance instruments for hospital characteristics.

Table 2.1 reports coefficients from pairwise regressions of demographic variable indicators, comorbidity indicators, and ambulance characteristics on each ambulance instrument, while controlling finely for patient location using zip code by origin fixed effects. The first

Table 2.1: *Balance of Patient Characteristics by Hospital Instruments*

	Black Patients		White Patients	
	Ambulance mean: Hospital Share Black (log)	Ambulance mean: Hospital Patient Volume (log)	Ambulance mean: Hospital Share Black (log)	Ambulance mean: Hospital Patient Volume (log)
Age 70-74	0.00236	-0.000778	0.00271*	0.00648**
Age 75-79	-0.00750*	0.00297	0.00226	0.0000926
Age 80-84	0.000899	-0.000570	-0.000117	-0.00113
Age 85-89	-0.00492	-0.00577	-0.00544***	0.00163
Age 90-94	0.00411	-0.00437	-0.00161	-0.00426
Age 95+	0.00141	0.00543	-0.000685	-0.00276
Gender: Male	0.00496	-0.0147	0.00477**	0.00333
Comorbidity: Hypertension	-0.00888*	0.0177	0.00152	0.00580
Comorbidity: Stroke	-0.000177	0.00669	0.000497	0.000109
Comorbidity: Cerebrovascular disease	0.00294	-0.00112	0.000322	0.000942
Comorbidity: Renal failure disease	-0.00576	0.0229**	0.000115	0.00458*
Comorbidity: Dialysis	-0.000617	0.00676*	0.000566*	0.00122**
Comorbidity: COPD	-0.00153	0.00437	0.00105	-0.000903
Comorbidity: Pneumonia	-0.00600*	0.0143**	-0.0000148	-0.00111
Comorbidity: Diabetes	-0.00620	0.0218**	0.000357	0.00383
Comorbidity: Protein calorie malnutrition	-0.00665*	0.00458	-0.0000880	0.00233
Comorbidity: Dementia	0.00140	0.00265	-0.00132	-0.00286
Comorbidity: Paralysis	-0.00437	0.00481	0.000699	0.000850
Comorbidity: Peripheral vascular disease	-0.00492	0.0125*	0.00110	0.00322
Comorbidity: Metastatic cancer	-0.00181	0.00615	0.000107	0.00174
Comorbidity: Trauma	-0.00468*	0.00291	0.000989	-0.000326
Comorbidity: Substance abuse	0.000965	0.00498	0.000235	-0.000791
Comorbidity: Major psychological disorder	-0.000928	-0.00826*	0.00119	-0.00168
Comorbidity: Chronic liver disease	-0.000467	0.00293*	0.000120	-0.0000899
Ambulance: Miles traveled with patient	-0.528***	-0.426	-0.166**	-0.341*
Ambulance: Advanced life support	-0.0291**	0.0771***	-0.0131*	0.0240*
Ambulance: Emergency traffic	0.00839	0.00874	-0.000254	0.00780
Ambulance: Payment	6.038	20.12	3.161	14.71*
Predicted 180-day mortality	-0.00319	0.00796**	-0.00111*	-0.000276
Predicted 360-day mortality	-0.00323	0.00843**	-0.00121*	-0.000577

Notes: Table shows balance across controls used in regressions. Estimates show the coefficient on the instrument in pairwise regressions of each characteristic on each instrument, with 3-digit ICD-9 principal diagnosis code fixed effects, year fixed effects, and zip code by origin fixed effects. The first two columns report results for Black patients (N = 85,425), and the last two columns report results for white patients (N = 627,296). Predicted 180-day and 360-day mortality are constructed by regressing observed 180-day and 360-day mortality on the full set of controls using OLS. * p<0.05, ** p<0.01, *** p<0.001.

two columns report results for Black patients, and the last two columns report results for white patients. Given the large number of regressions, we would expect some statistically significant associations between the ambulance instruments and patient characteristics. Indeed there are statistically significant coefficients in all columns – for both instruments and for both Black and white patients – but the magnitudes are small across the board. These results are consistent with the quasi-random assignment of ambulance companies, conditional on patient location.

The table does show some lack of balance among the ambulance variables. In particular, ambulances that take patients to hospitals with a higher Black patient share travel fewer miles on average and are less likely to offer advanced life support capabilities. Ambulances that take patients to higher volume hospitals also travel fewer miles on average but are more likely to offer advanced life support capabilities. These patterns are consistent across the Black and white patient samples, lending support to the differences-in-differences analysis discussed in Section 2.3.3.

Finally, as an aggregated test of balance along patient characteristics, I construct predicted 180-day and 360-day mortality using OLS and the full set of comorbitidies and other controls observed in the data. The associations between these predicted mortality measures and the ambulance instruments are small and contrast with the larger coefficients found in reduced form regressions on actual mortality.

The set of characteristics reported in Table 2.1 are included as control variables in the regression specifications for the main results, along with principal diagnosis code fixed effects, year fixed effects, and zip code by origin fixed effects. Throughout the results tables I report robustness and sensitivity to including the comorbidity and ambulance controls.

2.4.2 First stage

I show that ambulance assignment is associated with hospital assignment in Table 2.2. Assignment to an ambulance company that takes other patients to hospitals with a higher Black patient share is strongly linked to being treated at a hospital with a higher Black

Table 2.2: *First Stage by Patient Race*

	Share Black (log)			Patient Volume (log)		
	(1)	(2)	(3)	(4)	(5)	(6)
A. Black Patients						
Ambulance mean:	0.372***	0.371***	0.365***	0.00789	0.00815	0.0164
Share Black (log)	(0.0237)	(0.0237)	(0.0236)	(0.0132)	(0.0131)	(0.0132)
Ambulance mean:	-0.0310	-0.0306	-0.0361	0.604***	0.602***	0.610***
Patient Volume (log)	(0.0386)	(0.0386)	(0.0385)	(0.0259)	(0.0257)	(0.0256)
N	85425	85425	85425	85425	85425	85425
B. White Patients						
Ambulance mean:	0.531***	0.531***	0.532***	0.0129**	0.0129**	0.0138***
Share Black (log)	(0.0133)	(0.0133)	(0.0133)	(0.00399)	(0.00399)	(0.00404)
Ambulance mean:	0.0271	0.0271	0.0278	0.605***	0.605***	0.605***
Patient Volume (log)	(0.0163)	(0.0163)	(0.0163)	(0.0114)	(0.0114)	(0.0114)
N	627296	627296	627296	627296	627296	627296
Como. Xs		✓	✓		✓	✓
Amb. Xs			✓			✓

Notes: Table shows first stage results for the sample of Medicare ambulance transports between 2003 and 2014. Columns 1-3 report coefficients from the first stage regression of the ambulance instruments on the hospital characteristic share Black, sequentially adding in comorbidity controls and ambulance characteristic controls (Equation 2.3). Columns 4-6 report coefficients from the first stage regression of the ambulance instruments on hospital patient volume (Equation 2.4). All regressions include controls for 5-year age bins, sex, 3-digit ICD-9 principal diagnosis code, and year, as well as zip code by origin fixed effects. Panel A reports results for Black patients, and Panel B reports results for white patients. Standard errors clustered at the hospital service area (HSA) level are reported in parentheses. * p<0.05, ** p<0.01, *** p<0.001.

patient share; assignment to an ambulance company that takes other patients to hospitals with higher patient volume is strongly linked to being treated at a higher volume hospital. The first stage estimates are consistent with and without comorbitidy and ambulance controls and are highly statistically significant, for both Black (Panel A) and white (Panel B) patients.

The magnitudes of the first-stage coefficients range from 0.37 to 0.61. These are significantly less than 1, but are consistent with other ambulance-instrument work and the described natural experiment. When an ambulance company picks up a patient in a zip code outside of its primary service area, it is more likely to transport the patient back to its usual hospitals but not at the same rate as it transports the many patients living closest to its base of operations. In addition, the first-stage coefficient provides insight into the share of “compliers" in the IV natural experiment; some patients’ admitting hospitals are not affected by the preferences of the assigned ambulance company. This generates a strong positive correlation, but not one that is one-to-one.

For both Black and white paitents, the first-stage coefficient is smaller for the ambulance instrument predicting hospital share Black (0.37 to 0.53) than for the ambulance instrument predicting hospital patient volume (0.60 to 0.61). The first-stage coefficient is smaller still for Black patients than for white patients, for the hospital share Black instrument. The smaller first-stage coefficients indicate natural experiments with fewer IV compliers. Conversely, these are natural experiments with more “always takers" or “never takers," whose hospital admission is not affected by being picked up by an ambulance with a stronger affiliation to hospitals with a higher Black patient share.

Although the first-stage coefficients are strong from an IV estimation perspective for both instruments and patient samples, ambulance assignment appears to affect admitting hospital racial demographics less strongly than admitting hospital size. The corresponding 2SLS estimates will report the local average treatment effect (LATE) of hospital characteristics for the complier groups, with the caution that these complier groups may be different or represent different shares of the population by race and by characteristic.

2.4.3 Black Patient Share and Mortality: 2SLS Estimates

Building on the IV assumptions of exogeneity and relevance, Panel A of Table 2.3 shows that Black patients who are assigned to hospitals with a 10 percent higher Black patient share are 0.25 to 0.31 percentage points less likely to die in the 180 days following the hospital admission. This coefficient is robust across specifications with and without comorbitidy and ambulance controls, as shown in Columns 1-3.

Given that the specification includes zip code by origin fixed effects, the natural experiment compares patients who are quasi-randomly assigned among the set of local hospitals. The mean of the predicted hospital Black patient share for Black patients is 11.9 percent, with a within-zip code by origin standard deviation (SD) of predicted $log(\text{Black patient share})$ of 0.308. Therefore, the 2SLS estimates imply that a 1 SD (30.8 percent or 3.7 pp on the mean) increase in the hospital Black patient share among local hospitals reduces 180-day patient mortality by approximately 0.77-0.95 percentage points. This reduction represents a 2.5-3.1 percent decrease on a mean 180-day mortality of 30.8 percent.

The effect on 180-day mortality of being taken to a hospital with a higher share of patients who are Black is of a similar magnitude to the effect of being seen at a hospital with higher total patient volume. Indeed, using the same steps as above, being taken to a hospital with a 1 SD (15.7 percent) higher hospital volume among local hospitals reduces 180-day patient mortality by approximately 0.70 percentage points, using the strongest coefficient, from Column 3. However, the coefficient on Share Black fades and becomes not statistically significant for 360-day mortality, whereas the effect of being seen at a higher volume hospital persists strongly over time.

Given the results in Panel A, one possible conclusion would be that hospitals with a higher share of Black patients are of better quality in general for 180-day outcomes. However, I show in Panel B of Table 2.3 that for the white patient sample, being taken to a hospital with a higher share of patients who are Black has no effect on mortality 180 or 360 days following the hospital admission. Therefore, the mortality effect at 180 days is specific to Black patients.

Table 2.3: *2SLS Results by Patient Race*

	180-day mortality			360-day mortality		
	(1)	(2)	(3)	(4)	(5)	(6)
A. Black Patients						
Share Black (log)	-0.0314*	-0.0272*	-0.0254*	-0.0201	-0.0156	-0.0137
	(0.0128)	(0.0127)	(0.0127)	(0.0154)	(0.0151)	(0.0151)
Patient Volume (log)	-0.0294	-0.0395*	-0.0449**	-0.0260	-0.0379*	-0.0426**
	(0.0166)	(0.0162)	(0.0162)	(0.0168)	(0.0165)	(0.0163)
N	85425	85425	85425	85425	85425	85425
B. White Patients						
Share Black (log)	0.0043	0.0044	0.0048	0.0033	0.0033	0.0037
	(0.0030)	(0.0029)	(0.0029)	(0.0032)	(0.0030)	(0.0030)
Patient Volume (log)	-0.0167**	-0.0183***	-0.0192***	-0.0201***	-0.0219***	-0.0224***
	(0.0054)	(0.0052)	(0.0051)	(0.0055)	(0.0054)	(0.0054)
N	627296	627296	627296	627296	627296	627296
Como. Xs		✓	✓		✓	✓
Amb. Xs			✓			✓

Notes: Table shows two-stage least squares (2SLS) estimates of the relationship between hospital demographics and patient outcomes, for Medicare beneficiaries between 2003 and 2014. Columns 1-3 report results for mortality 180 days following hospital admission, and Columns 4-6 report results for mortality 360 days following the hospital admission, as in the model in Equation 2.5. For each outcome, comorbidity controls and ambulance controls are added sequentially. All regressions include controls for 5-year age bins, sex, 3-digit ICD-9 principal diagnosis code, and year, as well as zip code by origin fixed effects. Panel A reports results for Black patients, and Panel B reports results for white patients. Standard errors clustered at the hospital service area (HSA) level are reported in parentheses. * p<0.05, ** p<0.01, *** p<0.001.

Table 2.4: *2SLS Differences-in-differences Results*

	180-day mortality			360-day mortality		
	(1)	(2)	(3)	(4)	(5)	(6)
Share Black	-0.0341***	-0.0355***	-0.0348***	-0.0168	-0.0190	-0.0181
× Black	(0.0101)	(0.00983)	(0.00983)	(0.0121)	(0.0117)	(0.0116)
Patient Volume	-0.0141	-0.0231	-0.0267	-0.00737	-0.0180	-0.0214
× Black	(0.0145)	(0.0143)	(0.0143)	(0.0155)	(0.0152)	(0.0151)
Share Black	0.00399	0.00379	0.00424	0.00316	0.00286	0.00324
	(0.00295)	(0.00285)	(0.00282)	(0.00313)	(0.00301)	(0.00299)
Patient Volume	-0.0166**	-0.0182***	-0.0190***	-0.0202***	-0.0219***	-0.0224***
	(0.00530)	(0.00516)	(0.00509)	(0.00544)	(0.00537)	(0.00533)
N	712721	712721	712721	712721	712721	712721
Como. Xs		✓	✓		✓	✓
Amb. Xs			✓			✓

Notes: Table shows two-stage least squares (2SLS) differences-in-differences estimates of the relationship between hospital demographics and patient outcomes, for Medicare beneficiaries between 2003 and 2014. Columns 1-3 report results for mortality 180 days following hospital admission, and Columns 4-6 report results for mortality 360 days following the hospital admission, as in the model. in Equation 2.6. For each outcome, comorbidity controls and ambulance controls are added sequentially. All regressions include controls for 5-year age bins, sex, 3-digit ICD-9 principal diagnosis code, year, and the interaction between each control and an indicator for the patient being Black, as well as zip code by origin fixed effects. Standard errors clustered at the hospital service area (HSA) level are reported in parentheses. * p<0.05, ** p<0.01, *** p<0.001.

In Table 2.4, I further the comparison by estimating the differences-in-differences model described in Section 2.3.3, comparing the pattern for Black patients to that of white patients as a control group. Consistent with the previous results, I show that Black patients are 0.35 percentage points less likely to die in 180 days following the hospital admission when taken to a 10 percent higher share Black hospital, relative to white patients taken to a 10 percent higher share Black hospital. The differential effect for Black patients with regards to being taken to a higher volume hospital is not statistically significant, but the overall effect of being taken to a higher volume hospital is significant across specifications, as would be expected from prior work.

2.4.4 Interpretation and Robustness

In addition to our assessment of the exogeneity and relevance assumptions, interpretation of the coefficients on "Share Black" in Tables 2.3 and 2.4 as the causal effect of being treated at a hospital with a higher Black patient share requires some additional considerations. First, I note that the coefficients describe a standard linear approximation of the relationship between the hospital characteristic and patient outcomes. This approach indicates that, on average across the data observed, Black patients admitted to hospitals with a higher Black patient share have reduced 180-day mortality outcomes. However, the actual relationship may not be linear and may be driven by particular hospitals or parts of the distribution of hospital Black patient share.

Following this note, the coefficients for the white patient sample are similarly estimated as a standard linear approximation. Therefore, although I restrict to a set of hospitals and patient zip codes with common support across Black and white patient samples, the distribution of patients across these zip codes and hospitals may differ by race and affect the weighting underlying the linear approximations. However, the results are suggestive that hospitals may have race-specific causal effects and demonstrate that, within the variation observed in practice, Black patients benefit on average from being shifted towards hospitals with higher Black patient shares while white patients do not.

Second, I consider the implications of the differing first-stage coefficients documented in Table 2.2. In particular, because the first-stage coefficient for being admitted to a hospital with a higher Black patient share is smaller for Black patients than for white patients and smaller than the first-stage coefficient for being admitted to a higher volume hospital, the reduced form coefficient is scaled more than the others in the IV estimate. To address this concern, I report the reduced form results in Table 2.5, which display the same qualitative patterns as the 2SLS estimates and remain of a substantial magnitude.

Finally, I consider robustness of the results to mortality outcomes at different intervals following the hospital admission. Figure 2.2 plots reduced form coefficients and 95% confidence intervals using cumulative mortality in each month following the hospitalization as the outcome variable. We find that the largest reductions in mortality among Black patients are seen in months 4-6 following hospitalization, but the coefficients are consistently negative beginning in month 3. No such pattern is observed among white patients, and in fact the coefficients are consistently positive, implying increased mortality if anything at hospitals with a high Black patient share.

In addition, Figure 2.2 compares the reduced form coefficients, which are driven by variation in ambulance company assignment, with conditional OLS coefficients, which are estimated from regressions including the full set of controls including zip code by origin fixed effects but capture the relationship between Black patient share at the chosen hospital and patient mortality. The OLS results are strongly different from the reduced form pattern, with a generally positive correlation between hospital share Black and patient mortality among Black patients and no relationship for white patients. The contrast between the OLS and reduced form results highlights the importance of using quasi-random variation to draw conclusions about hospital effects.

2.5 Conclusion

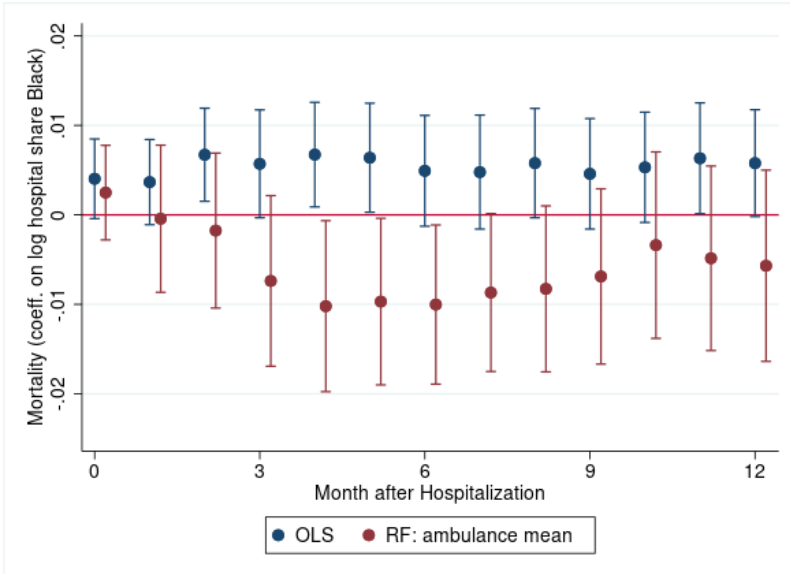
Using data on emergency Medicare patients within the years 2003 to 2014, I show that Black patients quasi-randomized to hospitals that treat more Black patients have lower 180-day

Table 2.5: *Reduced Form Results by Patient Race*

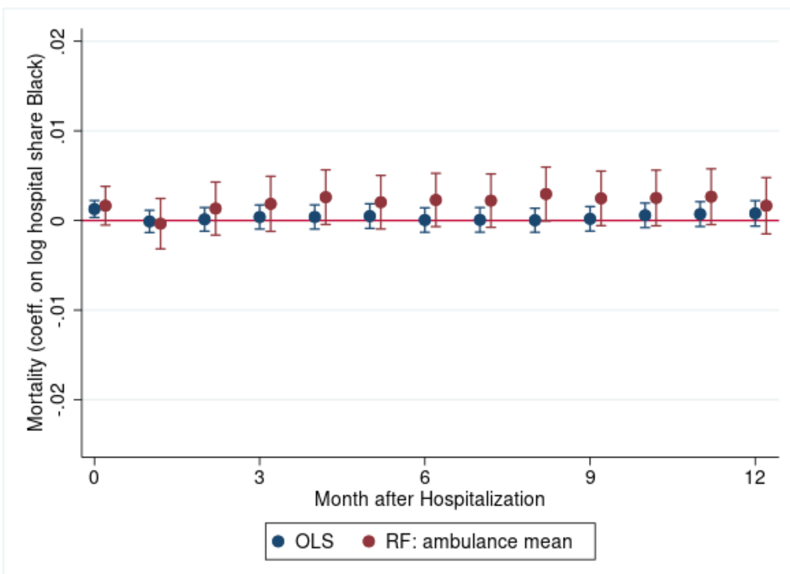
	180-day mortality			360-day mortality		
	(1)	(2)	(3)	(4)	(5)	(6)
A. Black Patients						
Ambulance mean:	-0.0119*	-0.0104*	-0.0100*	-0.0077	-0.0061	-0.0057
Share Black (log)	(0.0046)	(0.0046)	(0.0045)	(0.0056)	(0.0055)	(0.0055)
Ambulance mean:	-0.0168	-0.0230*	-0.0264**	-0.0151	-0.0223*	-0.0255**
Patient Volume (log)	(0.0099)	(0.0096)	(0.0096)	(0.0101)	(0.0098)	(0.0098)
N	85425	85425	85425	85425	85425	85425
B. White Patients						
Ambulance mean:	0.0021	0.0021	0.0023	0.0015	0.0015	0.0016
Share Black (log)	(0.0016)	(0.0015)	(0.0015)	(0.0017)	(0.0016)	(0.0016)
Ambulance mean:	-0.0100**	-0.0110***	-0.0115***	-0.0121***	-0.0131***	-0.0134***
Patient Volume (log)	(0.0032)	(0.0031)	(0.0031)	(0.0033)	(0.0033)	(0.0033)
N	627296	627296	627296	627296	627296	627296
Como. Xs		✓	✓		✓	✓
Amb. Xs			✓			✓

Notes: Table shows reduced form estimates of the relationship between ambulance propensities and patient outcomes, for Medicare beneficiaries between 2003 and 2014. Columns 1-3 report results for mortality 180 days following hospital admission, and Columns 4-6 report results for mortality 360 days following the hospital admission. For each outcome, comorbidity controls and ambulance controls are added sequentially. All regressions include controls for 5-year age bins, sex, 3-digit ICD-9 principal diagnosis code, and year, as well as zip code by origin fixed effects. Panel A reports results for Black patients, and Panel B reports results for white patients. Standard errors clustered at the hospital service area (HSA) level are reported in parentheses. * p<0.05, ** p<0.01, *** p<0.001.

Figure 2.2: *Mortality Following Hospitalization: Reduced Form*



(a) *Black Patients*



(b) *White Patients*

Notes: Figure shows coefficients and 95 percent confidence intervals from OLS (blue series) and reduced form (red series) regressions of hospital share Black on cumulative mortality outcomes one to twelve months after hospitalization for non-discretionary conditions. Panel A shows results for Black patients, and Panel B shows results for white patients. Sample includes admissions which arrived by ambulance transport to the emergency department between 2003 and 2014. All regressions include controls for 5-year age bins, sex, 3-digit ICD-9 principal diagnosis code, comorbidities, ambulance characteristics, year, and zip code by origin fixed effects. Standard errors are clustered at the hospital service area (HSA) level.

mortality. The results are robust to inclusion of controls for patient comorbidities and a set of ambulance characteristics that describe differences in pre-hospital care. The magnitude of the effect is large, with admission to a hospital with a 1 SD increase in Black patient share generating a similar reduction in mortality as admission to a hospital with a 1 SD increase in total patient volume. Further, the effect of being admitted to a hospital with a higher Black patient share is race-specific, with no corresponding effect for white patients.

This study provides evidence of race-specific hospital effects and the presence of positive Roy selection on heterogeneous hospital gains by race. Although the ambulance-instrument complier group for Black patients assigned to hospitals with higher Black patient shares is smaller than complier groups for other ambulance-instrument variations, the local average treatment effect (LATE) for those whose hospital choice *is* affected by the ambulance assignment is large. However, taking a step back from the 2SLS model, the reduced form coefficients also show substantial reductions in mortality and rely only on the assumption of exogeneity of the ambulance instrument.

This work indicates that further research into and support for hospitals, as well as ambulance companies, that serve Black patients as a large share of their total patient population may be important and meaningful for reducing racial disparities in health outcomes. The results underscore a need to consider heterogeneous treatment effects and nuanced models of patient outcomes in health care. The consideration of heterogeneous treatment effects adds to a growing literature arguing the importance of subgroup analysis and the creation of data and metrics by race and other subpopulations, especially for groups that are historically vulnerable and may be left behind. In the quest to improve health care efficiency and reduce costs, paying close attention to allocative efficiency, which has both distributional and overall welfare implications, may be key.

Chapter 3

The Black Physician Pipeline: Evidence from Medical School Records

Abstract

I show the limited progress in outcomes of Black medical school applicants over the years 1979-2020. Black applicants have grown to 10% of the applicant pool in recent years, but acceptance rates remain below those of white and Asian applicants. For Black students who do matriculate to medical school, graduation rates lag behind those of white and Asian medical students. Among the cohorts of graduated MDs, medical schools affiliated with historically Black colleges and universities (HBCUs) continue to play an important role in graduating Black physicians, accounting for 14.9% of Black physicians in 1984-1999 and 14.7% in 2000-2015. Taken together, the evidence highlights continuing gaps for Black students in the physician pipeline, with need for targeted actions.

3.1 Introduction

Black Americans make up 14.2% of the population but only 5.3% of active physicians (AAMC, 2022). This underrepresentation mirrors patterns in many high-income careers, including science, engineering, law, and management occupations (Wilson *et al.*, 2021). Given the history of oppression of African Americans in the U.S., the continued underrepresentation of Black Americans in high-status and high-income professions is a cause for examination. In patricular, it is important to critically examine the barriers that Black individuals face in such career pipelines in order to identify external factors which may contribute to underrepresentation today.

In this paper, I focus on the physician pipeline and document trends in the outcomes of Black medical school applicants, from 1979 through 2020. Using data on applicant, enrollee, and graduation cohorts by race/ethnicity from the Association of American Medical Colleges (AAMC), I find evidence of stagnancy in the graduation of Black doctors. Although the number of Black medical students has increased, particularly in recent years since 2016, acceptance rates and graduation rates for Black students remain below those of white and Asian peers and have stayed relatively constant over the past 40 years.

Despite the general stagnancy in acceptance rates and graduation rates, I show that there were periods of change. In the mid-1990s, at the height of advocacy for affirmative action, acceptance rates to medical school for Black applicants equaled and even surpassed those of white and Asian applicants. However by the mid-2000s the difference in acceptance rates re-emerged and persisted over the following decades.

Throughout these periods of change, the graduation rate differential between Black and white students remained constant. Therefore, although acceptance rates varied, there was no evidence of a change in the relative quality of admitted students by race. Black students who were admitted during the height of affirmative action were no less likely to graduate than Black students admitted in earlier or later periods.

However, the persistent Black-white graduation rate gap suggests that, without external forces, medical schools interested in maintaining high graduation rates may choose to admit fewer Black students. It follows that the schools that have been instrumental to producing Black physicians are those with health care diversity and Black education in their mission, including medical schools affiliated with Historically Black Colleges and Universities (HBCUs): Howard University College of Medicine, Meharry Medical College, and Morehouse School of Medicine. These three schools are just as important to the graduation of Black physicians today as they were in the 1980s.

This paper contributes to existing work on the role of medical schools in influencing physician diversity. In particular, the AAMC has advocated for and evaluated the success of efforts to improve minority representation in medicine over several decades. In 1991, it launched a 10-year plan with the goal of enrolling a class with 3,000 students from underrepresented minorities by the year 2000 (Cohen *et al.*, 2002). This initiative appeared to see immediate results but was hindered by attacks on affirmative action that began in the late-1990s (Garces and Mickey-Pabello, 2015; Mickey-Pabello and Garces, 2018). My results support this narrative of progress and reversal, while extending the time series to recent years and exploring connections to acceptance rates and graduation rates.

This work also relates to a growing literature on the importance of physician diversity from a public health perspective. Having a same-race doctor has been found to improve primary care and hospital outcomes for Black patients (Alsan *et al.*, 2019; Hill *et al.*, 2020). Black physicians are more likely to practice in underserved areas and communities with high percentages of Black residents and, controlling for the racial makeup of the community, care for significantly more Black patients, including more patients covered by Medicaid (Komaromy *et al.*, 1996; Laditka, 2004; Xierali and Nivet, 2018). More broadly, diversity in the health care workforce can improve the breadth and depth of the U.S. health research agenda and expand the pool of medically trained individuals who may take up leadership positions in the health care system.

3.2 Data and Methods

3.2.1 Data Source

The data consist of historical tabulations of applicants, matriculants, enrollees, and graduates of U.S. medical schools from the Association of American Medical Colleges (AAMC). In particular, the applicant and matriculant data tables contain annual counts of self-reported race/ethnicity among applicants, acceptees, and matriculants to U.S. medical schools, from application year 1978-1979 through 2019-2020. The enrollment and graduation data tables contain counts of self-reported race/ethnicity separately for each medical school and academic year, from academic year 1979-1980 through 2019-2020. The enrollment data originate from the AAMC Student Records System (SRS).

3.2.2 Measurement of Race

Information on race and ethnicity is collected from multiple AAMC sources including the Medical College Admission Test (MCAT), the American Medical College Application Service (AMCAS), and the Electronic Residency Application Service (ERAS). However, the methodology for collecting data on race/ethnicity has changed over time. Prior to academic year 2002-2003, an individual could identify with one and only one race/ethnicity category, whereas starting from academic year 2002-2003 individuals could select multiple race categories. In the analysis, I include individuals who have designated multiple races/ethnicities in each of the race/ethnicity categories that were selected.

In relevant figures, I include a dashed line to highlight the year after which individuals could designate multiple races.

3.2.3 Estimation of Graduation Rates

I construct approximate graduation rates. For each year t, race r, and medical school m, I observe student enrollment E across all four years, as well as the number of graduates G in the given year. I do not observe enrollment numbers for any specific graduating class. I therefore estimate graduation rates g as follows:

$$g_{t-1.5,r,m} = \frac{\sum_t^{t+3} G_{trm}}{E_{trm}}. \tag{3.1}$$

For example, for the 1984-1985 enrolled class, I compute the number of graduates observed in 1985, 1986, 1987, and 1988 as the numerator. I assign this graduation rate to year $t - 1.5 = 1982.5$, because it represents graduates from matriculation years 1981, 1982, 1983, and 1984. Broadly, this is similar to a 4-year graduation rate, but because I do not have enrollment counts for specific graduating classes some students could take more time and still be included as graduating.

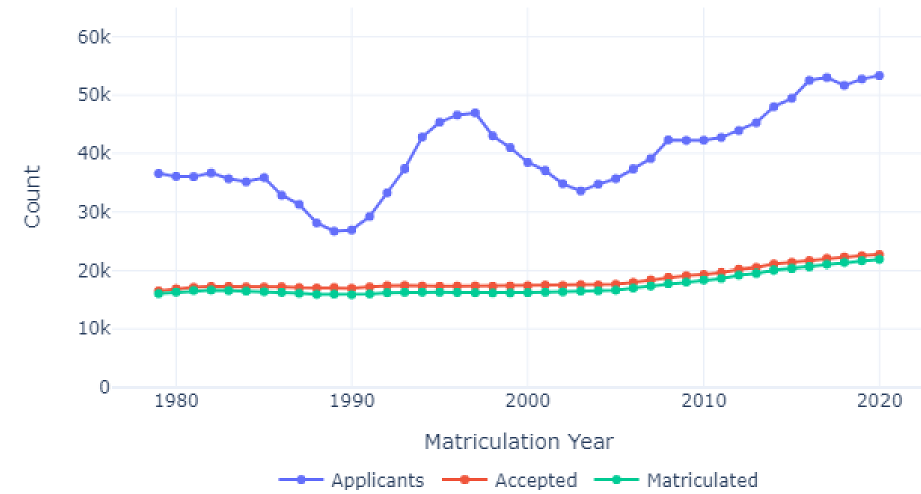
3.3 Results and Discussion

I begin by showing trends in the number of applicants to medical schools over the years 1979-2020. The number of applicants has fluctuated widely, reaching a low of 26,709 applicants in 1989, a peak of 46,965 applicants in 1997, and growing steadily since 2004 to a high of 53,371 applicants in 2020 (Figure 3.1a). The same qualitative patterns in applicant counts are present among Black applicants (Figure 3.1b), as well as across other observable subgroups, including first-time applicants, white applicants, and Asian applicants.

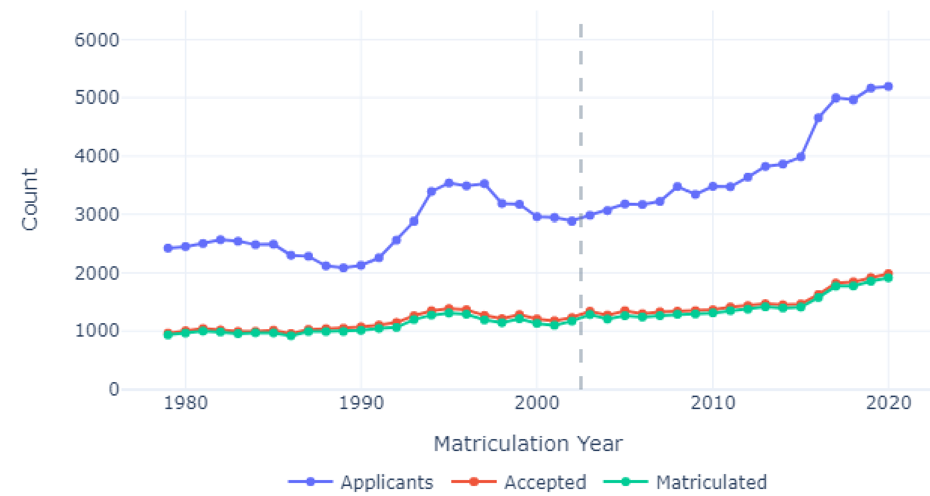
Despite ebbs and flows in the number of applicants, Figure 3.1a shows that the total number of accepted and matriculated medical students remained relatively consistent, before growing slowly since 2005. In contrast, there are a few noticeable patterns in the number of Black accepted and matriculated students over the observed time period (Figure 3.1b). In particular, the mid-1990s saw an increase in the number of Black accepted and matriculated students, from 992 matriculated students in 1989 to a peak of 1,309 matriculants in 1995. However, this increase was short-lived, decreasing back down to 1,131 matriculants by 2000 and remaining relatively steady for the following decade. Since then, there has been a notable increase in Black accepted and matriculated students starting in 2016, corresponding with a sharp increase in the number of applicants.

To contextualize the variation in the number of Black matriculants in terms of the racial

Figure 3.1: *Applicants, Acceptees, and Matriculants, 1979-2020*



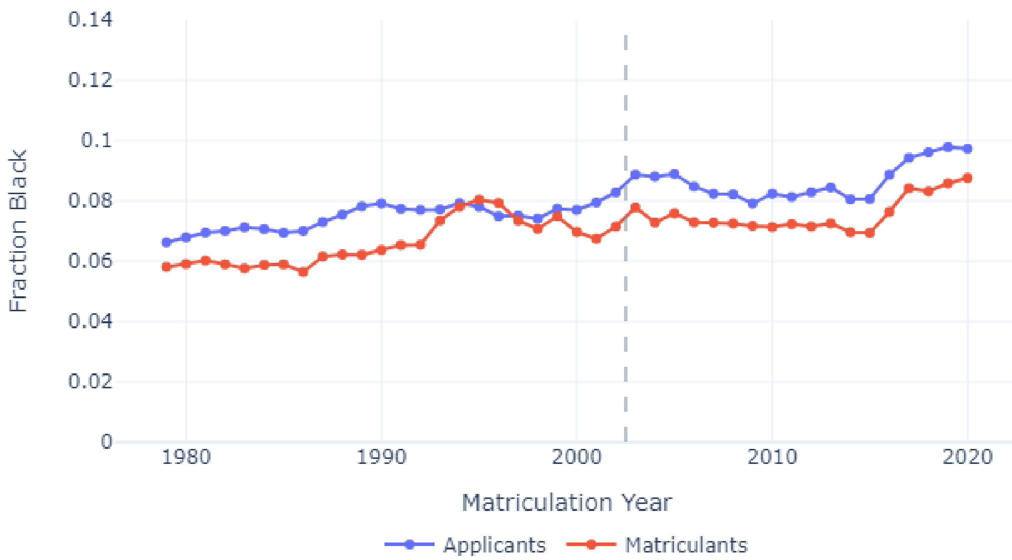
(a) *Total*



(b) *Black*

Notes: Data are from the AAMC Applicant and Matriculant Data File. Panel A shows the total number of applicants, acceptees, and matriculants to U.S. medical schools for academic years 1979-1980 through 2020-2021. Panel B shows the number of applicants, acceptees, and matriculants who self-identified as Black or African American. Prior to matriculation year 2003-2004, individuals could identify with one and only one race/ethnicity. The dashed line indicates the year after which individuals could designate multiple races.

Figure 3.2: *Black Share of Applicants and Matriculants, 1979-2020*



Notes: Data are from the AAMC Applicant and Matriculant Data File. The series show the number of applicants and matriculants to U.S. medical schools who self-identified as Black or African American divided by the total number of applicants and matriculants for each academic year, from 1979-1980 through 2020-2021. Prior to matriculation year 2003-2004, individuals could identify with one and only one race/ethnicity. The dashed line indicates the year after which individuals could designate multiple races.

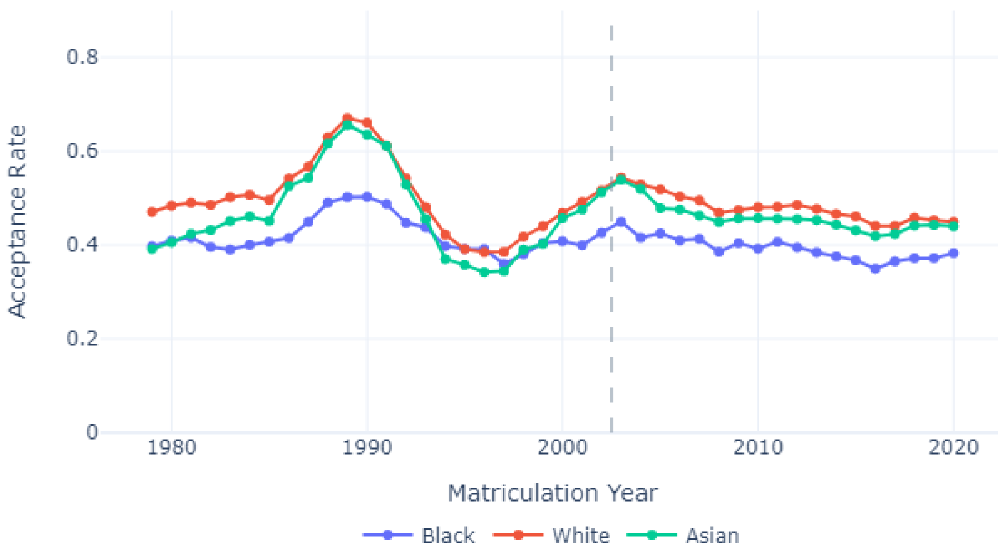
composition of medical students, Figure 3.2 shows the *share* of applicants and matriculants who identify as Black over the study period. We see that the Black applicant share has grown slowly over time, from 6.6 percent of the applicant pool in 1979 to 9.7 percent in 2020, but still remains below the population share of 14.2 percent in 2020 (Jones *et al.*, 2021). In addition, the Black matriculant share has consistently been below the Black applicant share, by about 1pp, with the exception of the mid-1990s, when Black students made up about 8 percent of both total applicants and matriculants.

From the perspective of an individual Black applicant, the convergence of Black applicant and matriculant shares in the mid-1990s suggests that acceptance rates may have increased for Black applicants during this period. However, Figure 3.3 shows that acceptance rates for Black applicants have remained relatively consistent over the study period, varying between a high of 50 percent in 1990 and a low of 35 percent in 2016. Acceptance rates for white and Asian medical school applicants have fluctuated more widely, between highs of 67 percent and 66 percent respectively in 1989 and lows of 38 percent and 34 percent respectively in 1996. In particular, acceptance rates for white and Asian applicants dipped in the mid-1990s to equal and fall below those of Black applicants. By the mid-2000s, acceptance rates for white and Asian students diverged from those of Black applicants again, remaining more than 5pp above acceptance rates for Black applicants in recent years.

The difference in acceptance rates could reflect overall differences in the quality of applicants, as admissions committees take into consideration many factors including GPAs, MCAT scores, volunteer work, and research when selecting applicants. To investigate whether differences in applicant quality may explain the persistent Black-white differential in acceptance rates, I plot an observable metric of student quality: graduation rates.

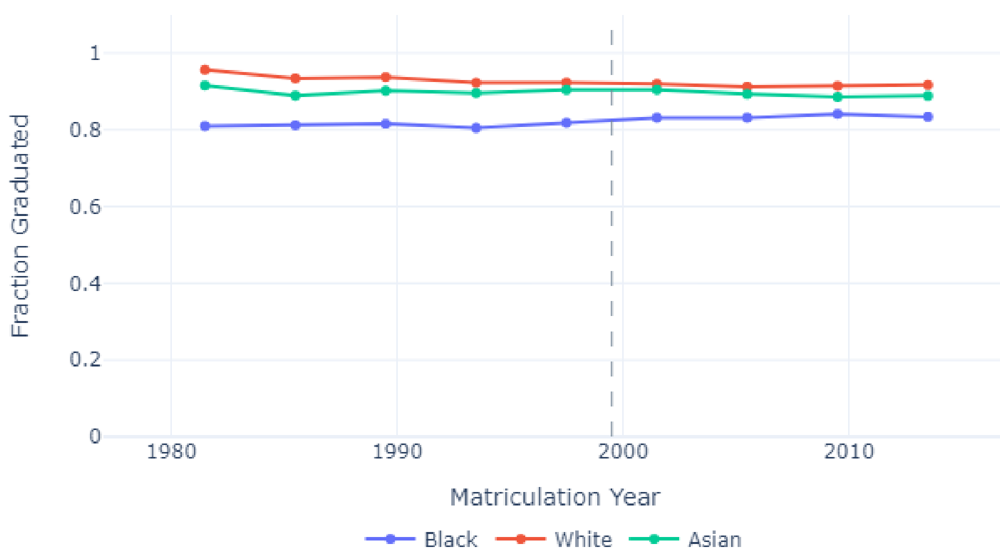
In Figure 3.4 I show that Black medical students across the study period had lower graduation rates than white and Asian students. However, I do not observe any differential trends in graduation rates for Black students relative to white and Asian students, while the acceptance rate differential widened, closed, and reemerged again. The graduation rate for Black medical students remained steadily around 81 percent in the 1980s and 1990s,

Figure 3.3: *Acceptance Rates by Race, 1979-2020*



Notes: Data are from the AAMC Applicant and Matriculant Data File. The series show the number of acceptees divided by the number of applicants to U.S. medical schools by race, for individuals who self-identified as Black or African American, White, and Asian. Prior to matriculation year 2003-2004, individuals could identify with one and only one race/ethnicity. The dashed line indicates the year after which individuals could designate multiple races.

Figure 3.4: *Graduation Rates by Race, 1980-2015*



Notes: Data are from the AAMC Student Records System (SRS). The series show estimated graduation rates as described in Section 3.2.3, for students who self-identified as Black or African American, White, and Asian. Prior to academic year 2002-2003, individuals could identify with one and only one race/ethnicity. The dashed line indicates the point after which individuals could designate multiple races. The dashed line occurs at an earlier matriculation year because the graduation rates are computed from data collected at a later date.

before increasing to 83 percent with the composition change once students could designate multiple races. The consistency of the Black-white graduation rate gap, as admissions policies varied, suggests that there is limited scope on the margin for selection of applicants who are more likely to be successful, within race.

On the other hand, the persistent Black-white graduation rate gap, which averaged 9.6 percentage points across years, does provide an explanation for lagging acceptance rates among Black applicants. If medical schools aim to improve their graduation rates, they may accept fewer Black students who may be statistically less likely to graduate. Further, the gap is not driven by sorting across schools: controlling for medical school the Black-white graduation rate gap remains at 10.4 percentage points.

Given the challenges in admitting and graduating Black medical students, I turn to examine the most successful educators of Black physicians during the study period. Table 3.1 shows that medical schools affiliated with Historically Black Colleges and Universities (HBCUs) have played and continue to play a very important role in educating Black physicians. Howard University College of Medicine in D.C., Meharry Medical College in Nashville, TN, and Morehouse School of Medicine in Atlanta, GA were among the top 5 medical schools producing Black MDs in graduating years 1984-1999, and became the top 3 in 2000-2015. Together, these schools graduated nearly 15% of Black physicians both in 1984-1999 and in 2000-2015, through student cohorts with very high Black student shares of 60-80 percent.

In addition, other schools that have graduated many Black MDs include the University of Illinois College of Medicine in Chicago, IL, Wayne State University School of Medicine in Detroit, MI, SUNY Downstate Medical Center College of Medicine in Brooklyn, NY, and the UNC School of Medicine in Chapel Hill, NC. In contrast to the HBCU-affiliated medical schools, these tend to be large medical schools with Black student shares ranging from 7-11 percent.

Table 3.1: *Top 10 Schools Graduating Black Physicians, 1984-1999 and 2000-2015*

		Black Graduates (1)	Share Black (2)	Share of Total Black Graduates (3)	Graduation Rate (4)
A. 1984 - 1999					
1	Howard University COM	998	0.666	0.068	0.845
2	Meharry Medical College	841	0.754	0.057	0.725
3	University of Illinois COM	429	0.093	0.029	0.737
4	Wayne State University SOM	379	0.094	0.026	0.791
5	Morehouse SOM	361	0.786	0.024	0.808
6	Lewis Katz SOM at Temple University	296	0.104	0.020	0.876
7	UNC at Chapel Hill SOM	290	0.096	0.020	0.842
8	UCLA David Geffen SOM	281	0.114	0.019	0.824
9	SUNY Downstate Medical Center COM	274	0.100	0.019	0.849
10	Rutgers New Jersey Medical School	264	0.102	0.018	0.825
B. 2000 - 2015					
1	Meharry Medical College	1052	0.659	0.059	0.842
2	Howard University COM	1042	0.777	0.058	0.777
3	Morehouse SOM	537	0.727	0.030	0.913
4	University of Illinois COM	443	0.094	0.025	0.805
5	University of Texas Medical Branch SOM	361	0.087	0.020	0.717
6	Wayne State University SOM	348	0.110	0.019	0.983
7	SUNY Downstate Medical Center COM	326	0.100	0.018	0.905
8	UNC at Chapel Hill SOM	289	0.115	0.016	0.778
9	Medical University of South Carolina COM	261	0.099	0.015	0.774
10	GWU SOM and Health Sciences	258	0.066	0.014	0.837

Notes: Data are from the AAMC Student Records System (SRS). Abbreviations are used in school names for College of Medicine (COM) and School of Medicine (SOM). Column (1) reports the cumulative number of graduates who self-identified as Black or African American. Column (2) reports the share of graduates from each school who identified as Black. Column (3) reports the share of total Black graduates from the period accounted for by a specific school. Column (4) reports the estimated graduation rate among Black enrollees at each school. Prior to academic year 2002-2003, individuals could identify with one and only one race/ethnicity. Aftwards, individuals could designate multiple races and are included as long as they self-identified as Black.

3.3.1 Implications for Practice and Policy

The results show that, despite ongoing discussions about physician diversity in research and policy spheres, there has been marked stagnancy in achieving meaningful increases in newly graduated physicians from underrepresented groups. In particular, I highlight the Black-white graduation rate gap and its implications for admissions as a potential roadblock in efforts to increase Black enrollment in and graduation from medical schools. In response, policy levers that improve graduation rates of Black medical students, such as ensuring social, academic, and financial support during medical school as well as maintaining partnerships with K-12 education systems and colleges to strengthen preparedness, could in turn lead to higher acceptance rates at the admissions stage. Future research should seek to understand barriers facing Black medical students that lead to lower graduation rates.

On the other hand, given recent research regarding public health benefits of Black physicians for marginalized Black communities, we may believe that the societal impact of training more Black doctors is important enough that greater admissions despite lower graduation rates could be rationalized for Black students. In this case, an external actor beyond individual medical school interests would be required. To align medical school incentives, we may want to develop an "impact factor" to measure and rank public health impacts of medical schools beyond standard metrics such as graduation rates. Because of the divergence of interests, HBCUs and other schools with health care equity and physician diversity in their mission remain crucial for the education of Black doctors, and need to be supported if we hope to develop a more diverse physician workforce.

3.3.2 Limitations

This study has several limitations. First, the analysis and conclusions are based on aggregate time trends, and do not control for individual characteristics of applicants such as GPA, MCAT scores, courses taken, and demographic characteristics beyond race. Therefore, the relationships between applicant counts, acceptance rates, and graduation rates are suggestive and should not be interpreted as causally linked. Unmeasured social and economic confounders that shift over time may affect all documented series. Although the correlational analysis in this study cannot demonstrate causal mechanisms, it is a step toward determining which factors are likely to contribute to underrepresentation of Black Americans among physicians.

Second, this paper focuses solely on outcomes for students between application and graduation from US medical schools granting MD degrees. I do not observe data on other health care professions and degrees, such as DO, physician assistant, and nursing programs, which Black students may opt into with or without applying to MD-granting medical schools. These other degrees may act as substitutes for or complements to the graduation of Black physicians with MD degrees and their presence can affect the diversity of the health care workforce from the patient's perspective. Nevertheless, of these degrees, MD credentials generate the highest incomes and decision-making power and are therefore important to study to improve diversity and equity at the highest levels.

Finally, I do not observe data on portions of the physician pipeline before and after medical school, including K-12 and college education, medical specialty decisions, residency, and professional practice following medical school. These inputs and decision points represent many additional opportunities for intervention and potential challenges that deeply affect the overall idea of reaching close to parity between Black population and Black physician shares. However, documenting disparities within the medical school application and graduation process is a step towards understanding challenges within the full pipeline.

3.4 Conclusion

I show stagnancy in acceptance rates, graduation rates, and number of graduates among Black medical school applicants over the years 1979-2020. These results highlight challenges in the pipeline for Black physicians and the continuing role of mission-driven institutions in training and graduating Black doctors.

References

AAMC (2022). *2021 State Physician Workforce Data Report*. Tech. rep., Association of American Medical Colleges, Washington, DC.

Abowd, J. M., Kramarz, F. and Margolis, D. N. (1999). High Wage Workers and High Wage Firms. *Econometrica*, **67** (2), 251–333.

Adler, N. E., Glymour, M. M. and Fielding, J. (2016). Addressing Social Determinants of Health and Health Inequalities. *JAMA*, **316** (16), 1641–1642.

Allcott, H., Diamond, R., Dubé, J.-P., Handbury, J., Rahkovsky, I. and Schnell, M. (2019). Food Deserts and the Causes of Nutritional Inequality. *The Quarterly Journal of Economics*, **134** (4), 1793–1844.

Alsan, M., Garrick, O. and Graziani, G. (2019). Does Diversity Matter for Health? Experimental Evidence from Oakland. *American Economic Review*, **109** (12), 4071–4111.

— and Wanamaker, M. (2018). Tuskegee and the Health of Black Men. *The Quarterly Journal of Economics*, **133** (1), 407–455.

Arias, E. and Xu, J. (2015). United States Life Tables, 2015. p. 64.

Bonhomme, S., Holzheu, K., Lamadon, T., Manresa, E., Mogstad, M. and Setzler, B. (2020). How Much Should we Trust Estimates of Firm Effects and Worker Sorting? p. 76.

Card, D., Cardoso, A. R. and Kline, P. (2016). Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women *. *The Quarterly Journal of Economics*, **131** (2), 633–686.

—, Heining, J. and Kline, P. (2013). Workplace Heterogeneity and the Rise of West German Wage Inequality. *The Quarterly Journal of Economics*, **128** (3), 967–1015.

Chan, D. C., Card, D. and Taylor, L. (2021). *Is There a VA Advantage? Evidence from Dually Eligible Veterans*. Working Paper.

Chandra, A., Kakani, P. and Sacarny, A. (2020). Hospital Allocation and Racial Disparities in Health Care. p. 47.

— and Staiger, D. O. (2007). Productivity Spillovers in Health Care: Evidence from the Treatment of Heart Attacks. *Journal of Political Economy*, **115** (1), 103–140.

— and — (2020). Identifying Sources of Inefficiency in Healthcare*. *The Quarterly Journal of Economics*, **135** (2), 785–843.

Chetty, R. and Hendren, N. (2017a). The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects. p. 73.

— and — (2017b). The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates. p. 95.

—, —, Kline, P. and Saez, E. (2014). Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States*. *The Quarterly Journal of Economics*, **129** (4), 1553–1623.

—, Stepner, M., Abraham, S., Lin, S., Scuderi, B., Turner, N., Bergeron, A. and Cutler, D. (2016). The Association Between Income and Life Expectancy in the United States, 2001-2014. *JAMA*, **315** (16), 1750–1766.

CMS (2021). *CMS Fast Facts*. Tech. rep., CMS.

Cohen, J. J., Gabriel, B. A. and Terrell, C. (2002). The Case For Diversity In The Health Care Workforce. *Health Affairs*, **21** (5), 90–102.

Colwell, R. L., Narayan, A. K. and Ross, A. B. (2022). Patient Race or Ethnicity and the Use of Diagnostic Imaging: A Systematic Review. *Journal of the American College of Radiology*, p. S1546144022001168.

Cooper, Z., Doyle, J., Graves, J. and Gruber, J. (2022). *Do Higher-Priced Hospitals Deliver Higher-Quality Care?* Tech. Rep. w29809, National Bureau of Economic Research, Cambridge, MA.

Cutler, D. M. and Lleras-Muney, A. (2010). Understanding differences in health behaviors by education. *Journal of Health Economics*, p. 28.

Deryugina, T. and Molitor, D. (2020). Does When You Die Depend on Where You Live? Evidence from Hurricane Katrina. p. 97.

Dimick, J., Ruhter, J., Sarrazin, M. V. and Birkmeyer, J. D. (2013). Black Patients More Likely Than Whites To Undergo Surgery At Low-Quality Hospitals In Segregated Regions. *Health Affairs*, **32** (6), 1046–1053.

Doyle, J., Graves, J. and Gruber, J. (2019). Evaluating Measures of Hospital Quality: Evidence from Ambulance Referral Patterns. *The Review of Economics and Statistics*, **101** (5), 841–852.

Doyle, J. J., Graves, J. A., Gruber, J. and Kleiner, S. A. (2015). Measuring Returns to Hospital Care: Evidence from Ambulance Referral Patterns. *Journal of Political Economy*, **123** (1), 170–214.

Finkelstein, A., Gentzkow, M. and Williams, H. (2016). Sources of Geographic Variation in Health Care: Evidence From Patient Migration. *The Quarterly Journal of Economics*, **131** (4), 1681–1726.

—, — and — (2018). What Drives Prescription Opioid Abuse? p. 55.

—, — and — (2019). *Place-Based Drivers of Mortality: Evidence from Migration*. Tech. Rep. w25975, National Bureau of Economic Research, Cambridge, MA.

Garces, L. M. and Mickey-Pabello, D. (2015). Racial Diversity in the Medical Profession: The Impact of Affirmative Action Bans on Underrepresented Student of Color Matriculation in Medical Schools. *The Journal of Higher Education*, **86** (2), 264–294.

Geruso, M. (2012). Black-White Disparities in Life Expectancy: How Much Can the Standard SES Variables Explain? *Demography*, **49** (2), 553–574.

Goodman, D., Brownlee, S., Chang, C.-H. and Fisher, E. (2010). *Regional and Racial Variation in Primary Care and the Quality of Care among Medicare Beneficiaries*. Tech. rep., Dartmouth Atlas Project.

Gray, D., Barton, B., Azam, I. and Bonnett, D. (2017). 2017 National Healthcare Quality and Disparities Report. p. 130.

Hanchate, A. D., Paasche-Orlow, M. K., Baker, W. E., Lin, M.-Y., Banerjee, S. and Feldman, J. (2019). Association of Race/Ethnicity With Emergency Department Destination of Emergency Medical Services Transport. *JAMA Network Open*, **2** (9), e1910816–e1910816.

Harper, S., Rushani, D. and Kaufman, J. S. (2012). Trends in the Black-White Life Expectancy Gap, 2003-2008. *JAMA*, **307** (21), 2257–2259, publisher: American Medical Association.

Hill, A., Jones, D. and Woodworth, L. (2020). *Physician-Patient Race-Match Reduces Patient Mortality*. SSRN Scholarly Paper ID 3211276, Social Science Research Network, Rochester, NY.

Hull, P. (2020). Estimating Hospital Quality with Quasi-Experimental Data. *SSRN Electronic Journal*.

Institute of Medicine (2003). *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care (with CD)*. Tech. rep., National Academies Press, Washington, D.C.

Jones, N., Marks, R., Ramirez, R. and Rios-vargas, M. (2021). *2020 Census Illuminates Racial and Ethnic Composition of the Country*. Tech. rep., US Census Bureau.

Kaiser Family Foundation (2002). *Racial/Ethnic Differences in Cardiac Care: The Weight of the Evidence*. Tech. rep.

Kline, P., Saggio, R. and Sølvsten, M. (2020). Leave-Out Estimation of Variance Components. *Econometrica*, **88** (5), 1859–1898.

Komaromy, M., Grumbach, K., Drake, M., Vranizan, K., Lurie, N., Keane, D. and Bindman, A. B. (1996). The Role of Black and Hispanic Physicians in Providing Health Care for Underserved Populations. *New England Journal of Medicine*, **334** (20), 1305–1310.

Laditka, J. N. (2004). Physician supply, physician diversity, and outcomes of primary health care for older persons in the United States. *Health & Place*, **10** (3), 231–244.

Lahiri, K. and Pulungan, Z. (2021). Racial/Ethnic Health Disparity in the U.S.: A Decomposition Analysis. p. 14.

Malek, S. K., Keys, B. J., Kumar, S., Milford, E. and Tullius, S. G. (2011). Racial and ethnic disparities in kidney transplantation: Disparities in kidney transplantation. *Transplant International*, **24** (5), 419–424.

Mickey-Pabello, D. and Garces, L. M. (2018). Addressing Racial Health Inequities: Understanding the Impact of Affirmative Action Bans on Applications and Admissions in Medical Schools. *American Journal of Education*, **125** (1), 79–108.

Molitor, D. (2018). The Evolution of Physician Practice Styles: Evidence from Cardiologist Migration. *American Economic Journal: Economic Policy*, **10** (1), 326–356.

Oaxaca, R. (1973). Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review*, **14** (3), 693–709, publisher: [Economics Department of the University of Pennsylvania, Wiley, Institute of Social and Economic Research, Osaka University].

Oaxaca, R. L. and Ransom, M. R. (1999). Identification in Detailed Wage Decompositions. *Review of Economics and Statistics*, **81** (1), 154–157.

Roy, A. D. (1951). Some Thoughts on the Distribution of Earnings. *Oxford Economic Papers*, **3**, 135–146.

Schore, J., Brown, R. and Lavin, B. (2004). Racial Disparities in Prescription Drug Use Among Dually Eligible Beneficiaries. *Health Care Financing Review*, **25** (2), 14.

Song, J., Price, D. J., Guvenen, F., Bloom, N. and von Wachter, T. (2019). Firming Up Inequality. *The Quarterly Journal of Economics*, **134** (1), 1–50.

Williams, D. R. and Sternthal, M. (2010). Understanding Racial-ethnic Disparities in Health: Sociological Contributions. *Journal of Health and Social Behavior*, **51** (1_suppl), S15–S27.

Wilson, V., Miller, E. and Kassa, M. (2021). *Racial representation in professional occupations*. Tech. rep., Economic Policy Institute.

Xierali, I. M. and Nivet, M. A. (2018). The Racial and Ethnic Composition and Distribution of Primary Care Physicians. *Journal of Health Care for the Poor and Underserved*, **29** (1), 556–570.