Published in final edited form as:

J Health Econ. 2015 January ; 39: 1–16. doi:10.1016/j.jhealeco.2014.09.003.

Racial Segregation and Quality of Care Disparity in US Nursing Homes

Momotazur Rahman, PhD and

Department of Health Services Policy and Practice, Brown University, Box G-S121(6),
Providence, RI 02912, Tel: 401-863-1275

Andrew D. Foster, PhD

Department of Economics and Health Services Policy and Practice, Brown University, 64
Waterman street, Providence, RI 02912, Tel: 401-863-2537

Momotazur Rahman: Momotazur_Rahman@brown.edu; Andrew D. Foster: afoster@brown.edu

Abstract

In this paper we examine the contributions of travel distance and preferences for racial homogeneity as sources of nursing home segregation and racial disparities in nursing home quality. We first theoretically characterize the distinctive implications of these mechanisms for nursing home racial segregation. We then use this model to structure an empirical analysis of nursing home sorting. We find little evidence of differential willingness to pay for quality by race among first-time nursing home entrants, but do find significant distance and race-based preference effects. Simulation exercises suggest that both effects contribute importantly to racial disparities in nursing home quality.

1. Introduction

Substantial racial disparities in health care utilization and health outcomes are of central concern in the design of US health care policy. Individuals of different races tend to receive care from different sets of providers; and minorities on average receive care from relatively lower quality providers and have worse health outcomes. The factors that have been posited to contribute to these disparities include differences in economic resources, insurance coverage, and preferences on the part of patients, providers, or referral agents.

An emerging strand of literature has focused on the role of geographical differences in the quality of providers in perpetuating racial disparities in health (Baicker, Buckles, and Chandra 2006; Baicker, Chandra, and Skinner 2005; Baicker et al. 2004; Chandra 2003; O'Connor et al. 1999; Welch et al. 1993). One common argument is that in the presence of residential segregation, blacks are more likely to live in poor neighborhoods where the standard of available care is low; consequently, they receive relatively lower quality care.

© 2014 Elsevier B.V. All rights reserved.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

However, the extent to which this actually translates into racial segregation and racial disparities in quality of care depends critically on how those making decisions trade off geographic distance with other attributes of a provider or facility. Differences in how people of different races value these factors can have major implications for the effects of policies to reduce disparities, but the data and estimation challenges that must be met in order to examine such tradeoffs are high.

The general pattern of disparities in health care is also evident in the context of nursing home care (Smith 1990, 1993). Indeed, in recent work, Smith et al (Smith et al. 2008; Smith et al. 2007) show that by some measures nursing homes are more segregated than residential neighborhoods in the US and that differences in measures of nursing home quality by race are both large and persistent. While many of the same mechanisms that drive other types of racial disparities in health also affect nursing homes, there are some aspects of nursing home care that make it a particularly fruitful setting in which to examine these issues. In particular, the nature of nursing home care tends to highlight the importance of race per se rather than unobserved differences in economic well-being or insurance coverage that happen to be correlated with race. A substantial fraction of nursing home residents are elderly and enter nursing homes following a hospital stay. Among this population and because of Medicare skilled nursing facility (SNF) care coverage rules, there are effectively no racial differences in payment status at the time of admission in the US. The extent of legacy or future payment effects (e.g., past experience with a provider at a time when payment status may have been different or differences that will emerge after the Medicare coverage runs out) can in part be controlled through a careful selection of patients.¹

Choices based on the attributes of fellow patients are also likely to be particularly salient in the context of nursing homes because of the residential aspect of care. Race not only distinguishes individuals physically, but reflects differences in culture and behavior.² We may expect individuals to prefer the company of people of the same race because of perceptions, whether accurate or not, that they share similar tastes and life experiences and will be treated with greater respect by members of the same race. When choosing a nursing home, patients may prefer to go to one where their friends and family reside or have stayed. Individuals may be more likely to have social experiences with others of their own race than with those of other races, which might translate into race-based preferences when selecting a nursing home. The hospital staff member that helps patients to choose nursing homes may suggest only placements in which he/she believes a patient will be among others of a similar background and may direct patients of different races to different nursing homes. Nursing home management may adopt recruitment or hiring strategies that target a particular race and/or may practice outright discrimination.

¹Indeed, Smith and coauthors report that segregation in nursing homes was little affected by the introduction of Medicare and Medicaid, which improved overall access to nursing homes by African-Americans.

²We deliberately blur the distinction between race and culture. In practice, our data includes information on race, but our conceptualization of race-based preferences, as defined below, is largely driven by aspects of culture: shared values, practices, or experiences. Self-reported race reflects at least in part cultural identification and, in any case, we do not have an empirical basis in which to make such a distinction. We will thus use the term race throughout the paper, but it should be understood in this broader context.

In this paper, we develop a simple theoretical framework for considering sources of racial segregation and racial disparities in quality among nursing homes. We highlight, in particular, three different sources of racial segregation. First, individuals may tend to choose nursing homes that are close to their own homes, and thus nursing homes may reflect patterns of racial segregation across neighborhoods. Second, individuals or those who advise them may deliberately choose nursing homes with higher proportions of people of similar race. Third, due to payer status, income, or taste differences, blacks and whites may have differences in the willingness or ability to pay for quality, resulting in a segmented market that breaks along racial lines.

Using a newly available data set that links residence and nursing home information at the individual level, we examine these three mechanisms. The distance-based effect aids in identification of racial preference effects by inducing a relationship between nursing home composition and the racial composition of proximate neighborhoods. We find little evidence of differential ability or willingness to pay for quality by race among new entrants, but do find significant evidence of both distance and race-based preference effects.

Using the estimated model, we then carry out a series of policy simulations. The first simulation exercise suggests that segregation and the quality gap would be much smaller in the absence of race-based preference. The second simulation indicates that an exogenous increase in the quality of care in the lowest-performing nursing homes would reduce both segregation and the gap in quality. The third simulation implies that an increase in nursing home quality affects the payer mix of a nursing home without much effect on the racial composition of patients.

2. Model

We develop a model both to provide basic insight into the qualitative implications of different types of nursing home sorting and to provide a framework for subsequent estimation. The model uses a standard discrete choice structure that is explicitly spatial, incorporates variation in residential composition and nursing home quality, and includes preference weights for different nursing home attributes. We characterize these weights as patient “preferences” in keeping with much of the economics literature on residential choice (Bayer and McMillan 2005; Bayer, McMillan, and Rueben 2004) and for expositional convenience, but it is important to note that, in fact, these weights reflect the behavior of three distinct agents: the patients, hospital staff who direct patients towards nursing homes, and nursing home managers. Not only is it difficult to find a compelling source of identification for these different preferences but, for some purposes, the distinction in terms of whose preferences are being estimated is of secondary importance—whether the allocation itself is sensitive to racial composition may be significant in terms of quality disparities regardless of whether this allocation is driven by patients, hospital staff, or nursing home managers. Obviously for other purposes, particularly an assessment of welfare effects or in terms of targeting public policy, the distinction could be critical.

Our model also abstracts from possible interactions between short- and long-stay patients. This distinction allows us to solve formally for an equilibrium racial composition and is well

matched with our empirical focus on new entrants, but disregards that most nursing homes have both types of patients. Most patients enter nursing homes for Medicare-paid rehabilitation skilled nursing facility (SNF) care after acute hospitalization events. Medicare covers up to 100 days of SNF care, and the average SNF length of stay was 27 days (Medpac 2012). About 20% of SNF patients remain at nursing homes after Medicare coverage ends, pay out-of-pocket or by Medicaid or privately purchased supplemental insurance, and are considered long-stay (Intrator, Zinn, and Mor 2004; Lau et al. 2005). Currently, about 16% of nursing home residents are short-stay Medicare-paid SNF patients.

Prior research has suggested that nursing homes that are primarily Medicaid-financed are lower quality on average (Carter and Porell 2003; Gertler 1992; Mor et al. 2004; Stevenson 2006). This financing can lead to differences in quality by race to the extent that there are important racial differences in income and wealth among long-stay patients. Even if there is no difference in payment status by race for Medicare-financed short-stay patients, these racial differences in payment status for long-stay patients coupled with distance- or race-based preferences can lead to racial disparities in quality among short-stay patients. Rather than explicitly modeling this process of quality formation and the link between short- and long-stay patients, we capture this long-stay effect by taking nursing home quality as given and, consistent with the data, correlated with neighborhood racial composition. We discuss the implications of this simplification for estimation and for our counterfactual simulation exercises below.

Consider a population of nursing home patients consisting of two races, whites ($r=0$) and blacks ($r=1$), that is distributed exogenously across a series of J residential neighborhoods and who make choices over K nursing homes. The residential neighborhoods are discrete points along a line with neighborhood j located at the point x_j and having a fraction black r_j . Nursing homes are also discrete points along the line, with nursing home k located at point y_k . The endogenously determined fraction of blacks in nursing home k is denoted by n_k . Nursing home quality net of any out-of-pocket differences associated with quality is denoted by q_k . A preference for quality may thus be thought of as a willingness to pay for it.

The utility of individual i from neighborhood j if he enters nursing home k is

$$V_{ijk} = U_{ijk} + \varepsilon_{ijk} \quad (1)$$

where

$$U_{ijk} = -\beta_1(x_j - y_k)^2 - \beta_2(r_i - n_k)^2 + \beta_3 q_k + \beta_4 q_k r_i \quad (2)$$

and ε_{ijk} is an idiosyncratic taste that is shock specific. That is, the utility of a particular assignment decreases in the squared distance between the residential neighborhood (NB) location and the nursing home (N) location and in the squared difference between the individual's race and the racial composition of the nursing home, and increases in quality with a possible differential effect by race.

From a partial equilibrium perspective, assuming ε_{ijk} has a type-one extreme value distribution, the model reflects a standard discrete choice framework (McFadden 1974, 1978). Individuals take as given the nursing home composition and thus the probability that i in neighborhood j enters nursing home k is

$$P_{ijk} = \frac{\exp(U_{ijk})}{\sum_{j'} \exp(U_{ij'k})}. \quad (3)$$

Summing across all individuals' neighborhoods, the fraction of patients who are black in nursing home k is

$$\bar{n}_k = (\sum_i \sum_j P_{ijk} r_i) / \sum_i \sum_j P_{ijk}. \quad (4)$$

The average fraction black of nursing homes $\bar{n}_{rj}^*$, average neighborhood to nursing home (NB-NH) distance $\bar{d}_{rj}^*$, and average quality $\bar{q}_{rj}^*$ by origin neighborhood and race can be calculated by averaging the relevant quantity across nursing homes. For given parameters – nursing home and neighborhood locations, and nursing home quality – solving for an equilibrium involves finding a fixed point so that the nursing home racial composition (4) is consistent with individual patient choice (3).

We now turn to some simulations that illustrate the differential implications of race versus distance-based sorting, abstracting, for these purposes, from the presence of sorting by quality ($\beta_3 = \beta_4 = 0$). We assume that the J neighborhoods are evenly spaced in the unit interval $[0,1]$ (i.e., $x_j = (j-1)/(J-1)$) and that the K nursing homes are evenly spaced in a broader interval $[-1/4, 5/4]$, with $J=10$ and $K=16$.³ Further, we assume that the neighborhood fraction black rises linearly from zero to one as one moves from left to right so that $r_j = x_j$. In keeping with the empirical results reported below, nursing home quality is assumed to decline linearly from left to right, from $3/2$ for $k=1$ to 0 for $k=K$. We assume that the population of nursing home entrants is large and of equal size across neighborhoods and that nursing homes are never capacity constrained. Finally we consider two sets of parameter values, which differ in terms of the relative importance of distance-versus race-based preferences. In the distance-based scenario, $\beta_1 = 2$, $\beta_2 = 0.1$, while in the preference-based scenario, $\beta_1 = 0.1$, $\beta_2 = 2$.⁴

Figure 1 presents the simulated racial composition of nursing homes $\bar{n}_k$ by nursing home location and illustrates the basic point that the presence of segregated neighborhoods can be a consequence either of distance- or preference-based sorting. Note that, overall, the degree of segregation is roughly similar in the two cases, with the fraction of black residents

³Nursing homes are placed in a wider interval than neighborhoods in order to reduce truncation effects arising from the fact that individuals from neighborhoods at the ends of the interval have relatively fewer “close” nursing homes.

⁴Note that in principle we could set the race parameter to zero in the distance-based scenario; however, setting the distance parameter to zero in the preference-based scenario creates a multiplicity of equilibria since one might get highly segregated nursing homes but there would be no particular mechanism to place black nursing homes in black neighborhoods and white nursing homes in white neighborhoods.

ranging from 25% to 75%. The primary difference is the extent to which neighborhood sorting exhibits a kind of tipping (Card, Mas, and Rothstein 2008) in neighborhoods with roughly 50% black. In particular, nursing homes just to the right of this level become primarily black and those just to the left become primarily white. In contrast, with distance-based preference, nursing home composition follows proportionately the composition of the neighborhoods.

Figure 2 exhibits sorting from the perspective of individual patients. In particular, we plot P_{ijk} for black and white individuals from predominantly black neighborhoods ($r_j^- = .75$) by nursing home destination for the two preference scenarios. Note that the distributions in each case are relatively wide, reflecting the relative importance of idiosyncratic taste shocks that have been incorporated into the model given the parameter values and the assumed distribution. There are, nonetheless, clearly distinctive patterns by race and type of preference. In the distance-based scenario, the mode probability involves entering the nursing home geographically nearest that neighborhood, and there is very little difference between blacks and whites in nursing home composition conditional on neighborhood of residence. Conversely, in the race-based scenario, the racial composition of selected nursing homes of blacks and whites from the same neighborhood are quite different. Even given the importance of the idiosyncratic tastes evident in the distance-based distributions, there is a strong tendency for blacks (whites) to move towards nursing homes in more heavily black (white) neighborhoods.

While it is evident from Figure 1 that distinguishing between distance- and race-based preferences is difficult using just data on the racial composition of nursing homes in different neighborhoods, Figure 3, which plots distance d_{rj}^* , shows that there are important differences in distance traveled by blacks and whites by racial neighborhood composition for the two scenarios. In particular, with distance-based preferences, there is little difference between whites and blacks, and there is also little difference in the distance traveled by neighborhood of origin. By contrast, with race-based preferences, blacks in relatively white neighborhoods tend to travel relatively large distances, as do whites in relatively black neighborhoods.

Figure 4, which plots n_{rj}^* , shows that there are also differences in nursing home composition for the two scenarios. In particular, with distance-based preferences, the average fraction black in the nursing home rises linearly with neighborhood composition and does not differ appreciably by race. Conversely, under race-based preferences, the average racial composition of nursing homes for blacks and whites is relatively constant across neighborhoods of origin with the average fraction black being substantially higher for blacks from any neighborhood than for whites from any neighborhood.

The quality graph (Figure 5) is essentially the reverse of the racial composition graph (Figure 4).⁵ Under distance-based preferences, quality falls with neighborhood composition but is essentially the same for blacks and whites. Of course, given the higher fraction of blacks in the primarily black neighborhoods, this means that blacks on average receive lower quality care. For race-based preferences, there is little variation with neighborhood composition, but whites receive comparably higher levels of care, with the difference being

about 20% of the difference between the highest and lowest quality nursing homes in the simulation.

Addition of a preference for quality that is the same by race ($\beta_3 > 0$, $\beta_4 = 0$) would in general moderate these effects. For example, if, as assumed, quality is inversely correlated with fraction black in the neighborhood, in the distance-based scenario, both blacks and white in relatively black neighborhoods would travel farther, and segregation by neighborhood of origin would diminish. In the preference-based model, blacks in relatively black neighborhoods would travel farther, but whites would not travel as far; again segregation by neighborhood of origin would diminish. If there were racial differences in preference so that blacks, for example, had lower willingness to pay for quality ($\beta_4 < 0$), then segregation could be enhanced by variation in quality and this could enhance disparities in care.

3. Data

These methods put extreme demands on the data. This paper uses individual level-data from four different sources: Medicare enrollment records, Medicaid Analytical eXtract (MAX), Medicare part A (for skilled nursing facility and hospital) claims, and the Minimum Data Set (MDS). Zip code level data comes from Census 2000 aggregates.⁶ We have used the Certification and Survey Provider Enhanced Reporting (CASPER) (formerly known as Online Survey, Certification, and Reporting [OSCAR] data) for nursing home attributes and American Hospital Association (AHA) data for hospital attributes. More detail on the data sets is provided in Appendix A.

Our analysis is restricted to the Medicare fee-for-service⁷ beneficiaries who entered a nursing home for Medicare SNF care between July 1, 2004, and June 30, 2005, after a hospitalization and do not have any history of nursing home residence since 1999 based on Medicare claims and MDS data. There are several reasons to take this subset of patients. First, a key assumption of our analysis is that residential location before going to a nursing home is exogenous and not affected by the choice of nursing home.⁸ We use as the residential address the location two years before entry into a nursing home. Second, as noted, these patients are more homogenous in terms of insurance coverage compared to repeat, long-term, or returning patients. We use only short-stay patients who are transferred from general hospitals, as specialized hospitals may have distinctive racial compositions and/or transfer policies.

We find 810,362 fee-for-service Medicare beneficiaries discharged from general hospitals to SNFs between July 1, 2004, and June 30, 2005, who were not prior nursing home residents.

⁵The values of betas 3 and 4 in Figure 5 are zero. If quality varies by neighborhood composition, then even in the absence of quality preference differences, there will be substantial differences in the quality of nursing homes, and these differences will be influenced by the relative importance of distance and race-based preferences. Adding a quality preference that is the same for the races will have little impact on this graph. If whites have a higher preference for quality, however, this would tend to produce something that looks like race-based preferences in Figure 5.

⁶There are about 42,000 zip codes in the entire United States, but Census aggregate variables are available for 32,023 zip codes.

⁷There are two types of Medicare enrollees: fee-for-service and managed care (about 14% in 2004–5). Medicare claims for the beneficiaries enrolled in managed care are not available and are thus excluded from the analysis.

⁸While those who have been previously housed in a nursing home may have subsequently adjusted their residential location to be close to that nursing home, we expect that the attributes of a proximate nursing home are a relatively minor consideration in terms of residential choice for patients who have not previously entered a nursing home.

We drop individuals for whom SNF, hospital, and residential zip code identification did not match with the CASPER, AHA, and census data, respectively. We also exclude beneficiaries with race other than black and white and not residing in the contiguous 48 states, leaving us with 646,494 beneficiaries in the study population. Of these individuals, 7.8% are black.

As we are interested in the sorting behavior of patients of different races, it is important to note the differences in other attributes like insurance status and stay length, because of the possibility that sorting is driven by other attributes that correlate with race. Nursing home specialization in particular medical conditions could lead to the appearance of race-based preferences, for example, if the incidence of these conditions is different by race. While we are looking at nursing home choice at entry and thus variation of payment status among our study population at the time of entry is minimal, we recognize that entering patients may anticipate a longer stay, at which point payment status may become a significant factor.

Table 1 presents the descriptive statistics of the individual characteristics that have been used in this paper. Individuals' demographics have been obtained from Medicare enrollment records. Educational attainment comes from MDS assessments. Medicaid enrollment status at the time of nursing home admission is obtained from the MAX. All the clinical variables are obtained from the Medicare claims of the hospital event prior to nursing home admission. Patients' residential zip codes are obtained from Medicare enrollment file. As shown in Table 1, there is also substantial dissimilarity in the demographic characteristics of patients of different races. Black patients are about two years younger and one percent more likely to be male and have higher comorbidities and lower education. Compared to the white patients, black patients reside in predominantly black, urban, and low-income neighborhoods.

Table 2 presents selected characteristics of destination nursing homes. To calculate the distance variables,⁹ we used exact coordinates of the nursing homes and hospitals obtained using the street addresses from the CASPER and AHA data respectively. For beneficiaries' residential locations, we used zip code centroids based upon the zip code tabulation area (ZCTA) definition file. The racial compositions of the nursing homes have been calculated from the MDS data for only the long-stay residents. The remaining nursing home variables are obtained from CASPER.

Characteristics of the destination nursing homes are quite different by race. On average, the proportion of black patients in those nursing homes where white patients were admitted is roughly seven percent, whereas the existing proportion of black patients in those nursing homes where black patients were admitted is 35 percent. Though the number of beds and

⁹To calculate the distance we have used the Haversine formula, which considers the distance between two points on a sphere from their latitude and longitude. It gives much better estimates than conventional spherical law of cosine especially in the case of short distances. The formula is:

$$A = (\sin(\Delta\text{lat}/2))^2 + \cos(\text{lat}1) * \cos(\text{lat}2) * (\sin(\Delta\text{long}/2))^2$$

$$C = 2 * \text{atan2}(A^{0.5}, (1-A)^{0.5})$$

$$\text{Distance} = C * 6371$$

occupancy rate is higher in destination nursing homes for black patients, the number of registered nurses per bed is lower.¹⁰

4. Descriptive Analysis

To undertake a preliminary test of the predictions of the theory developed above, we selected a population of roughly 360,000 individuals from 7,474 zip codes that have at least one nursing home patient of each race. Census-based data were used to determine the racial composition of the neighborhoods of residence based on zip codes. We first present polynomial regression plots of the relationship between nursing home composition and neighborhood racial composition and between nursing home distance and neighborhood racial composition by individuals' race. These relationships are plotted in Figures 6–8. Note that because the distribution of neighborhoods by percentage black is highly skewed (a large proportion of neighborhoods have very few blacks), we have used the log of percentage black in the zip codes on the horizontal in order to capture graphically the full variation in the data.

Figure 6 shows the graph of the distance¹¹ to nursing home in kilometers by racial composition of the neighborhood and corresponds in content to the simulation results presented in Figure 3. This graph seems to coincide with the predictions of the race-based theory. In particular, blacks living in predominantly white neighborhoods tend to travel farther than do whites in those same neighborhoods, and whites in predominantly black neighborhoods travel longer distances than blacks in the same neighborhoods.

Figure 7 shows the graph of the racial composition of nursing homes by racial composition of the neighborhoods of residence prior to nursing home entry and corresponds to Figure 4 from the simulations. Interestingly, it has features of both scenarios. On the one hand, both lines slope upwards as predicted by the distance-based theory. On the other hand, the lines are separated and roughly parallel as predicted by the theory of race-based segregation. Evidently, and perhaps not surprisingly, both distance- and race-based preferences seem to be factors in nursing home choice.

Quality, as measured by the adjusted number of nursing staff per bed¹² (Rahman et al. 2014a), also shows evidence of both distance- and race-based preferences (Figure 8). Overall, there is a decline in nursing home quality with neighborhood racial composition as predicted by the distance-based model except among those neighborhoods with the highest fraction black. However, as predicted by the race-based simulation model (that incorporates a quality gradient with respect to neighborhood racial composition), there is a gap between the quality of care for black and white nursing home entrants from a given neighborhood.

¹⁰Registered nurses (RN) are more skilled in terms of scope of practice, education, and clinical training compared to Licensed practical nurses (LPN). Certified nursing assistants (CNAs) provide bedside care under the supervision of an RN or LPN. In 2009, average salary is \$74k for RNs, \$49k for LPNs, and \$31k for CNAs.

¹¹Here we calculated median distance traveled by patients of different races in any zip code.

¹²Nursing staff per bed ratios were calculated using two variables: full time equivalent registered nurse (RN) per bed and full time equivalent licensed practical nurse (LPN) per bed, obtained from the OSCAR. Given that RNs and LPNs are not strictly comparable, we calculated an RN care equivalent of staffing per bed that is RN per bed plus LPN per bed multiplied by the wage ratio of LPN to RN (0.624). Thus it is in effect the RN per bed that could be purchased if the entire nursing budget were spent on RNs.

Overall, then, there is evidence of both distance- and raced-based preferences, and the process of sorting has an adverse impact on quality for black nursing home entrants. We now turn to a structural analysis, which provides the opportunity to simultaneously control for a variety of different characteristics that may affect nursing home choice and will provide a basis for a series of counterfactual simulations below.

5. Structural analysis

The first stage of structurally modeling nursing home choice behavior is to select an appropriate choice set (CS). Ideally, a choice set for an individual should reflect all the opportunities from which he/she is making a choice.¹³ Frequently, the empirical literature on discrete choices with product-specific alternatives (for example, choosing a commute to the office) assumes that individuals choose from the universe of alternatives. But in cases where individuals can choose from a large number of alternatives (for example, residential location choice, where each housing unit can be an alternative), researchers have used a randomly chosen sample of available alternatives following McFadden (1978).

In our case, the universal choice set is composed of about sixteen thousand nursing homes in the United States. We could have followed McFadden (1978) by taking a random subset from the universe set and assigning it to the individual patients. This approach might not be suitable for our case, however, because relevant alternatives would be randomly dropped and irrelevant alternatives (e.g., those that are very far away) would be randomly added.¹⁴ For example, if we were to take a random five percent of nursing homes to construct the choice set for an individual, we would have a choice set of roughly 850 different nursing homes. We would then have to drop off all the individuals who are not going to one of the randomly selected homes. If patients were evenly distributed across the nursing homes, we would then, on average have only five percent of patients left for our analysis. We used instead the ordered availability assumption of Ben-Akiva and Boccara (1995) to keep the set of alternatives small. In particular, we followed Rahman et al. (2013; 2014b) to construct three different residential zip-code specific choice sets and estimated our choice function using each of them. The first choice set, CS_{it}^1 is composed of any destination nursing home for any patient from a given zip code. The second choice set, CS_{it}^2 , is the union of the nursing homes within a 22-kilometer radius from the origin zip code and the geographically nearest 15 nursing homes from the origin zip code. The third choice set, CS_{it}^3 is the union of CS_{it}^1 and CS_{it}^2 . We also further restrict choice set three by limiting it to nursing homes in the state of residence. Details of the choice set are provided in Appendix B.

Our primary interest in this paper is in the racial composition and distance parameters. One reason the estimated distance parameter might be biased is that a patient may anticipate his need for nursing home residency and choose a residential location near to his preferred home prior to nursing home entry. It is for this reason that we took the residence zip code of

¹³For example, a nursing home that treats only patients with a certain diagnosis should not be an alternative for a person who does not have such diagnosis. However, it should not reflect the deterministic part of the choice rule; for example, a person's dislike of low-quality NH does not imply that low-quality nursing homes are not an option for him.

¹⁴For example, we may randomly drop a nursing home located in the same zip code where a patient lives from that patient's choice set while adding a very distant nursing in a different state where the patient is really unlikely to go.

the patient from the Medicare enrollment record in 2002 to measure the neighborhood-to-nursing home distance. A more subtle concern is that nursing homes in the long run partly determine their locations and quality with reference to potential future patients. For the purposes of this paper, we take nursing home location and quality as given. It is unclear whether or how this may bias our estimated choice parameters, but regardless, it is evident that our counterfactual simulations will not allow for adjustment along these dimensions. We will not, therefore, for example, be able to evaluate the effects of a policy change such as the elimination of certain types of nursing homes that operate through a change in the quality or location of other nursing homes. To address this limitation, we would require a detailed model of nursing home entry, closure, and quality choice, which is beyond the scope of this paper.

We do, however, consider racial composition to be endogenous. Because, in the presence of race-based preferences, patients make their choices in part based on nursing home racial composition, and these choices in turn affect the composition of the nursing home, we expect a multinomial logit model to overestimate the preference for racial homogeneity. There could be an omitted variable bias as well. Black patients may tend to cluster in particular nursing homes based on other unmeasured characteristics such as the racial composition of staff, the food that is served, or the wealth of other patients that happens to correlate with race. To partially deal with these issues of endogeneity, we use an instrumental variable strategy.

Finding an instrument for the racial composition of nursing home residents would seem to be a formidable task. However, assuming distance-based preferences and to the extent that we can take as given the geographical distribution of the population and the location of nursing homes, there is a potential source of identification that arises from the catchment area of a particular nursing home. In particular, the decision of people from one zip code to enter a nursing home in a given location will only be affected by the racial composition of other zip codes that are proximate to that nursing home through the racial composition of that nursing home.

Our previous linear model is not particularly helpful here because of the perfect correlation we imposed between racial composition and location. We thus consider a two-dimension version of the model but restrict attention to an illustrative case in which there are four equally sized neighborhoods surrounding two nursing homes. Figure 9 provides a graphical illustration of the structure we have in mind. In the case that there are no returns to nursing home attributes, the probability of any patient going to either nursing home given our model should be the same and equal to $1/2$. That is, regardless of the degree of segregation across neighborhoods, the same fraction of individuals in each nursing home should be black. If there is a return to proximity but this return does not differ by race, the racial composition of nursing homes will, in general, be different. In particular, while neighborhoods 1 and 4 will be evenly split between the two nursing homes, neighborhood 2 will primarily go to nursing home “A” and neighborhood 3 will primarily go to nursing home “B.” Thus, an increase in the proportion of residents black in neighborhood 3 raises the proportion black in nursing home B but has no impact on the entry of blacks from neighborhoods 1 and 4 into nursing home B. Finally, consider what might happen in the case of a preference for both racial

homogeneity and proximity. Suppose again that the share of blacks in neighborhood 3 exceeds that in neighborhood 2. Then, given a small preference for proximity, blacks will be differentially concentrated in nursing home B, and thus in neighborhoods 1 and 4, a higher share of blacks than whites will enter nursing home B. More generally, an increase in the share of blacks in neighborhood 3 will increase the entry of blacks from neighborhood 1, for example, to nursing home B, and decrease the entry of whites from that neighborhood to that home.

The exclusion restrictions here have a base plausibility but are not ironclad. Individuals may be concerned about travelling to a nursing home in a location that is dominated by people from a different race or income strata even if the nursing home itself houses people that are similar to them. Alternatively, a patient of a given race may be more likely to have a family member living in a neighborhood that is dominated by his own race than by some other race. If visitation from family members is a primary rationale for choosing a nursing home, then there could also be a relationship between nursing home choice and the racial composition of the neighborhood surrounding a nursing home net of the composition of the population in the nursing home.

For an estimation of the discrete choice model, we selected a random 20 percent (120,566 individuals) from the population that entered a nursing home for post-acute care. Table 3 provides a comparison of the results from the different choice sets. In Column 1, where we use the choice set based on revealed preference alone, 531 patients are eliminated, mainly because the number of available alternatives is one. In this case, many relevant alternatives are absent in the choice set and overall predictive ability is low. In the second column, where we use a proximity-based choice set, we had to drop about 10 percent of the individuals because the nursing homes they entered are not in their ascribed choice set; however, this approach seems to work well in other respects, with the goodness of fit increasing to 0.28 compared to 0.2 in Column 1. The difference in the distance parameter is also noticeable, suggesting it may be overestimated in Column 1 (where relatively distant options are not in the choice set) and underestimated in Column 2 (where the individuals who traveled higher distances are dropped out). In Column 3, we present the result of using the union of the previous two choice sets. Here the problem of exclusion of individuals is solved. In Column 4, we have used the choice set of Column 3 without the out-of-state nursing homes and we cluster the errors at the state level. In this case, we drop about 1.8% of the patients who were admitted to out-of-state nursing homes. This additional constraint does not affect the parameters but, as one might expect, the clustering has large effects on the estimated t-ratios. For the rest of the analysis we use CS_{it}^3 with the same-state nursing home restriction.

In Table 4, we present the parameter estimates of the choice equation with alternative specifications.¹⁵ The first column shows the results when only distance and the share of black residents in a nursing home determine the choice decision. The distance coefficients are negative for individuals of both races and statistically not different. As expected given

¹⁵Note that rather than estimating the race and distance parameters from equation (2) directly, we estimate the interaction term that is implied by expanding the quadratic term and incorporating the absolute value of distance.

the evidence of racial preferences, the coefficient corresponding to the share of black residents in a nursing home is negative for whites and positive for blacks. The second column shows the result with nursing input variables after controlling for size of the facility, occupancy rate, and a measure of acuity of the treated patients. Column 3 incorporates all the variables from the previous two columns. This is the specification that we will use for the rest of the paper. This model does not include factors like the share of Medicaid patients or profit status, which may be correlated with staffing quality. We want nursing variables to pick up the effects of these observed and unobserved differences in the quality of care between nursing facilities. Inclusion of observed markers of quality of care reduces the magnitude of the staffing variables and does not markedly affect the race or distance coefficients (comparing Columns 4 and 5).

To correct for the potential bias in the race coefficient, we apply a two-step control function approach (Petrin and Train 2010). In the first step, the endogenous variable is regressed on other explanatory variables and the instruments using OLS, and the residual is saved. In the second step, the saved residual is added as an additional regressor in the multinomial logit estimation. The first stage is reported in Appendix Table 1. Overall, the instruments explain 68% of the variation in nursing home share.

The second-stage results appear in Column 5 of Table 4. The race preference coefficients became bigger: the coefficient of share of black residents for white residents becomes more negative, and for black residents it becomes more positive. This may reflect a negative correlation between fraction black and unobserved nursing home attributes that are attractive to blacks but not whites. An alternative explanation is measurement error—it may well be that the racial composition in a particular year is a noisy estimate of the average racial composition over a longer time frame that actually governs choice.

Translation of these coefficients into marginal effects suggests that the consequences of distance- and race-based sorting for nursing home composition are substantial. We followed Rahman et al. (2014b) to calculate marginal effects following conditional logit estimation. If the distance from a randomly selected nursing home in the choice set increases by 10 km, holding all else constant, the probability of going to that nursing home declines by about 2 percentage points for both type of patients. If the share of blacks in a randomly selected nursing home increases by 10 points, the probability that a white patient will go to that nursing home declines by about 1.3 percentage points and that of a black patient increases by 0.2 percentage points. If the share of blacks in the nearest nursing home from a residential zip code increases by 10 points, the probability that a white patient will go to that nursing home declines by 6 percentage points and that of a black patient increases by 1 percentage point (see Appendix Figure 1).

For the sake of confirming robustness and to check that the race preference parameter is not driven by individual attributes that are correlated with race, we also estimate the choice function for different groups of individuals based on Medicaid eligibility, primary diagnosis, and the likelihood of becoming long-stay nursing home residents using Specification 4 in Table 5. The results are presented in Appendix Table 2. Distance and race preference parameters are very similar and statistically significant for all groups. Though all individuals

in our analysis are admitted to nursing homes for Medicare-paid post-acute care, individuals with different Medicaid eligibility have different nursing home choice rules. However, race preference is statistically different between black and white patients within groups of patients with the same Medicaid eligibility. Such differences are also practically the same for groups of patients with the same primary diagnosis from the hospital stay. Another factor that may drive differential race preference is a difference in length of stay at a nursing home between races. We defined an individual as long-stay if the length of stay at a nursing home exceeds 20 days. We chose this cutoff because the cost sharing for the first 20 days is zero and the following 80 days involve a copayment. 70% of the black and 60% of the white patients stayed at nursing homes for more than 20 days. We used a linear probability model to predict the likelihood of becoming a long-stay resident based on individuals' clinical, demographic, and neighborhood characteristics and estimated the choice function separately for individuals who belong to different quartiles of this likelihood. These estimated choice functions are presented in Appendix Table 3, and race parameters are virtually same for all quartile groups.

Given that hospitals play a key role in SNF selection process (Rahman et al. 2013), racially segregated hospitals (Dimick et al. 2013) may contribute to the observed segregation patterns in nursing homes. A key variable in our model is the distance of a nursing home from the originating hospital, which may partly capture location-based segregation of hospitals. We also estimated separate choice functions by rehospitalization rate and share of black patients (see Appendix Table 4). Though the share of black is much higher in hospitals with high rehospitalization rates, estimated race parameters are roughly the same. We also found that race parameters are larger for hospitals with low shares of blacks than hospitals with higher shares of blacks.

In addition, by using data from 48 states, our analyses include counties where the percentages of black residents are relatively small. This may result in the inclusion of observations in the data where a predominately black nursing home is not an option or is a relatively rare option. Hence, the empirical results could be driven by the construction of the choice set. Since the African American population is more concentrated in the southern region of the country, we estimated separate regressions for the southern region and the rest of the country and for metro and non-metro areas. We found similar results for different regions.

6. Simulations

One key advantage of discrete choice modeling is that we can predict the probability of choosing alternatives under different scenarios. In this section we conduct counterfactual simulation exercises that illustrate the importance of distance- and race-based preferences for racial disparities. In our first counterfactual simulation, we show the equilibrium sorting outcomes in a world in which racial composition in a nursing home has no effect on nursing home admission. This exercise is particularly important because it gives us a hint as to how much of the segregation and quality gap can be attributed to the evident preference for racial homogeneity that our estimates uncover. The objective of our second simulation is to get some idea of what policies would be most effective in reducing the racial gap in quality of

care. We compare two different interventions: one targets nursing homes with the lowest availability of RNs, and the other targets nursing homes with the highest share of black residents.¹⁶ In a third simulation, we examine how the composition of admitted patients in terms of race and Medicaid eligibility changes in response to an extra RN in a facility. It should be emphasized again that all simulations take the current quality and location of nursing homes and patient residences as givens.

The basic approach to solving for a new equilibrium consists of the following steps: first, based on the existing nursing home characteristics and preference parameters, the probability of each individual going to a particular nursing home in his choice set is predicted. Second, the probabilities of entry are adjusted so that they correctly predict the number of new admissions actually observed in that nursing home¹⁷ and sum to one across the choice set for each individual. This balancing is done in order to emphasize the potential of re-sorting to increase the levels and reduce disparities in care, and to downplay the possibility of (unrealistically) improving care by expanding the size of the most attractive nursing home at the same quality and location. To implement this correction, we deflate probabilities for each nursing home by the ratio of total probability assigned to the total number of admissions. We then divide the new probabilities for each individual by the sum of the new probabilities for that individual across his/her choice set to obtain a new set of individual probabilities. This process is then repeated until convergence is achieved. Third, based on the newly obtained probabilities, the racial composition variables are calculated for each nursing home. We then replace the old composition variables with the new ones in the preference equations. We again repeat this process until the composition variables converge.

As noted, one key assumption for this entire procedure is that nursing quality attributes such as the availability of nursing staff, the number of beds, and the number of competing facilities are not affected by the change in the composition variables. This assumption may not be as restrictive as it seems. As our analysis is focused on new patients, all of whom are covered by Medicare at entry, policy changes that primarily affect new entrants may have relatively little impact on long-stay patients, which may importantly drive nursing home decisions about staffing and overall quality because of differences in payer status.

6.1 If there were no race-based preference

We have already shown that patients have a strong tendency to go to nursing homes with a higher share of residents of their own race, for which they are evidently sacrificing quality of care or proximity if necessary. Here we consider the implications of sorting in the absence of race-based preferences. This exercise allows us to shed some light on the unexplained gap in the quality of care. As black patients or the agents that make nursing home choices for them evidently trade off quality of care for a higher share of black residents in a nursing home, we expect the quality gap to decline in the absence of race-based preference.

¹⁶We understand that this is a very racially insensitive intervention. However, this exercise has been done for theoretical interest, to check the difference in the outcomes.

¹⁷This is equivalent to a “market clearing condition” (like Bayer and McMillan [2005]), under the assumption that NHs are always fully occupied and they cannot adjust their capacities in the presence of higher demand.

Figure 10 presents the result as deviations in the hypothetical scenario from the benchmarks with regard to distance, racial composition, and quality. The upper panel shows the percentage change in distance traveled by both races. Under the new scenario, white residents from predominantly white neighborhoods would have traveled higher distances, while the opposite is true for the white residents in predominantly black neighborhoods. On the other hand, black residents in white neighborhoods are now traveling shorter distances to go to nursing homes with a higher proportion of black residents, and those in black neighborhoods are now travelling longer distances, perhaps for higher quality of care. The middle panel also shows the expected result that the equilibrium share of black residents in the destination nursing homes for white patients is much higher, whereas the opposite is true for black patients. Numerical results are presented in Table 5. Column 1 shows that, based on the predicted probabilities using estimated parameters, the share of black residents in the destination nursing homes for black patients is about 29% higher than for whites. When we control for residential zip code fixed effects, this figure is 14%. The results in the case of “no race preference” are presented in Column 3. There is still segregation, and the share of black residents where the black patients enter is about 21% higher. However, this gap is entirely due to residential segregation, as is evident in the fact that the gap disappears when we control for residential zip codes.

The graphs with nursing staff are shown in the lower panel of Figure 10. In the absence of race-based preference, black patients would have gone to nursing homes with a higher RN to total nurse ratio. The white patients would have gone to nursing homes with a lower RN to total nurse ratio, particularly those white residents of predominantly black neighborhoods, because they are now trading proximity for quality. As shown in Table 5, the quality gap vanishes in the case of no race preference. Evidently, a substantial part of the quality gap is due to the willingness of the patients or their agents to sacrifice quality of care for racial homogeneity.

6.2 Simulation of segregation and quality gap under alternative policies

Here we consider the counterfactual scenario in which a state provides a few selected nursing homes with an extra registered nurse, while all else remains unchanged. We consider two alternative ways of selecting the nursing homes: the lowest 10th percentile of nursing homes in terms of the RN to total nurse ratio and the top 10th percentile of nursing homes in terms of share of black patients. Thus, out of 15,693 nursing homes in our study, 1,578 are selected only for the staff-based intervention, and 1,556 are selected only for the racial composition-based intervention, and 313 firms get both interventions. Appendix Figure 1 shows that the target nursing homes by racial composition are mostly located in predominantly black neighborhoods while the distribution of treated nursing homes by the nursing staff is more even.

Both the interventions have a similar effect overall, but the race-based intervention, not surprisingly, yields a greater reduction in racial quality disparities. As shown in Table 6, both of the interventions increase segregation, i.e., the racial gap in the share of black residents in the destination nursing homes. The quality gap reduces immediately after an intervention, especially in the case of race-based intervention, and then increases again after

an adjustment of the racial composition of the competing nursing facilities. Changes in new equilibrium from before admission can be also seen from Figures 11 and 12. Most of the intervened facilities, as well as most of changes, are located in predominantly black neighborhoods. Blacks are going to facilities with higher share black, and whites are going to nursing homes with lower share black, contributing to an overall increase in segregation. Patients of both races enjoy higher quality of care after the intervention, and this increase is higher for black patients in the case of race-based intervention.

6.3 Does an increase in quality change NH patient composition?

In this section, we examine how a facility's composition of patients, especially in terms of race and Medicaid eligibility, changes in response to an exogenous change in quality. We estimate, in particular, a choice function interacted for four groups of patients stratified by race and Medicaid eligibility. The results are shown in Columns 1 and 2 of Appendix Table 2. While there is no evident difference in the return to quality by race within either the Medicaid or the non-Medicaid subgroups, the effect of quality on nursing home choice is substantially stronger for the latter group suggesting, as shown elsewhere, that higher quality nursing homes attract a higher share of non-Medicaid patients. To evaluate the magnitude of these effects, we chose 10% of the nursing homes at random and added an extra RN to each selected home and then allowed patient allocations to adjust. The simulations indicate that this increase in quality would increase the number of patients admitted by 2.4% in the selected nursing homes. The average percentage of blacks in these homes would decline by 0.1, while the average percentage of Medicaid eligibles would decrease by 0.7. But significantly, the increase in the number of Medicaid non-eligible patients and the reduction in the number of Medicaid eligible patients are predicted to be the same for both races.

7. Conclusion

We motivated our paper by noting the high degree of segregation in the nursing home industry as well as the discrepancy in the quality of care attained by patients of different races. We stress that while differences in residential locations and socioeconomic status help to explain these factors, racial sorting may also play an important role. We then presented a framework that considers how patients and/or their agents choose nursing homes and focus on the distinction between distance and race-based preferences. We used this framework to structure an empirical analysis of newly available data. We found that blacks from predominantly white neighborhoods travel a long way to nursing homes with a higher share of black residents and lower quality of care. In an effort to quantify the relative importance of different variables in nursing home choice, we estimated the choice process formally, defining three alternative choice sets for each zip code based on the available data. As the share of blacks in a given nursing home is, in the context of the model, determined endogenously, we used a control function approach to identify the racial preference effect, employing as instruments the racial composition of zip codes that are proximate to a given nursing home. There is strong evidence of race-based sorting, but the distance and quality preference parameters are roughly the same for both races. Using the estimated preference parameters, we conducted three counterfactual simulations, taking as givens the location and quality of existing nursing homes. The simulations suggest that policy interventions that

change the process of sorting across nursing homes can have substantial impacts on the levels and disparities in care.

From a methodological perspective, our paper also contributes to a growing literature on the spatial dimension of racial inequality and population distribution more generally. Among the innovations are a simple theoretical framework for considering distance versus race-based sorting, an integration of several different types of data in order to obtain relatively complete linked information on source and destination, a plausible source of identification for a model in which individual choices are simultaneously determined with the choices of their neighbors, and a series of counterfactual simulations that help to isolate the consequences of changes in sorting from more direct effects of policy changes on the quality of health care received.

References

- Baicker K, Buckles KS, Chandra A. Geographic variation in the appropriate use of cesarean delivery. *Health Aff (Millwood)*. 2006; 25(5):w355–67. [PubMed: 16895942]
- Baicker K, Chandra A, Skinner JS. Geographic variation in health care and the problem of measuring racial disparities. *Perspect Biol Med*. 2005; 48(1 Suppl):S42–53. [PubMed: 15842086]
- Baicker K, Chandra A, Skinner JS, Wennberg JE. Who you are and where you live: how race and geography affect the treatment of medicare beneficiaries. *Health Aff (Millwood)*. 2004; (Suppl Variation):VAR33–44. [PubMed: 15471775]
- Bayer P, McMillan R. Racial Sorting and Neighborhood Quality. NBER Working Paper. 2005; (11813)
- Bayer P, McMillan R, Rueben KS. What drives racial segregation? New evidence using Census microdata. *Journal of Urban Economics*. 2004; 56(3):514–35.
- Ben-Akiva M, Boccara B. Discrete choice models with latent choice sets. *International Journal of Research in Marketing*. 1995; 12(1):9–24.
- Card D, Mas A, Rothstein J. Tipping and the dynamics of segregation. *Quarterly Journal of Economics*. 2008; 123(1):177–218.
- Carter MW, Porell FW. Variations in hospitalization rates among nursing home residents: the role of facility and market attributes. *The Gerontologist*. 2003; 43(2):175–91. [PubMed: 12677075]
- Chandra AS, Skinner J. Geography and Racial Health Disparities. NBER Working Paper. 2003
- Dimick J, Ruhter J, Sarrazin MV, Birkmeyer JD. Black patients more likely than whites to undergo surgery at low-quality hospitals in segregated regions. *Health Aff (Millwood)*. 2013; 32(6):1046–53. [PubMed: 23733978]
- Gertler PJ. Medicaid and the Cost of Improving Access to Nursing-Home Care. *Review of Economics and Statistics*. 1992; 74(2):338–45.
- Intrator O, Zinn J, Mor V. Nursing home characteristics and potentially preventable hospitalizations of long-stay residents. *J Am Geriatr Soc*. 2004; 52(10):1730–6. [PubMed: 15450053]
- Lau DT, Kasper JD, Potter DE, Lyles A, Bennett RG. Hospitalization and death associated with potentially inappropriate medication prescriptions among elderly nursing home residents. *Arch Intern Med*. 2005; 165(1):68–74. [PubMed: 15642877]
- McFadden, D. *Frontiers in Econometrics*. New York: Academic Press; 1974. Conditional Logit Analysis of Qualitative Choice Behavior; p. 105–42.
- McFadden, D. Modeling the choice of residential location. In: Karlqvist, A.; Lundqvist, L.; Snickars, F.; Weibull, J., editors. *Spatial Interaction Theory and Planning Models*. North-Holland; Amsterdam: 1978. p. 75–96.
- Medpac. *A Data Book: Health Care Spending and the Medicare Program*. 2012
- Mor V, Zinn J, Angelelli J, Teno JM, Miller SC. Driven to tiers: socioeconomic and racial disparities in the quality of nursing home care. *Milbank Q*. 2004; 82(2):227–56. [PubMed: 15225329]

- O'Connor GT, Quinton HB, Traven ND, Ramunno LD, Dodds TA, Marciniak TA, Wennberg JE. Geographic variation in the treatment of acute myocardial infarction: the Cooperative Cardiovascular Project. *JAMA*. 1999; 281(7):627–33. [PubMed: 10029124]
- Petrin A, Train K. A Control Function Approach to Endogeneity in Consumer Choice Models. *Journal of Marketing Research*. 2010; 47(1):3–13.
- Rahman M, Foster AD, Grabowski DC, Zinn JS, Mor V. Effect of hospital-SNF referral linkages on rehospitalization. *Health Serv Res*. 2013; 48(6 Pt 1):1898–919. [PubMed: 24134773]
- Rahman M, Gozalo P, Tyler D, Grabowski DC, Trivedi A, Mor V. Dual Eligibility, Selection of Skilled Nursing Facility, and Length of Medicare Paid Postacute Stay. *Med Care Res Rev*. 2014a
- Rahman M, Grabowski DC, Gozalo PL, Thomas KS, Mor V. Are dual eligibles admitted to poorer quality skilled nursing facilities? *Health Serv Res*. 2014b; 49(3):798–817. [PubMed: 24354695]
- Smith DB. Population ecology and the racial integration of hospitals and nursing homes in the United States. *Milbank Q*. 1990; 68(4):561–96. [PubMed: 2292991]
- Smith DB. The racial integration of health facilities. *J Health Polit Policy Law*. 1993; 18(4):851–69. [PubMed: 8120348]
- Smith DB, Feng Z, Fennell ML, Zinn J, Mor V. Racial disparities in access to long-term care: the illusive pursuit of equity. *J Health Polit Policy Law*. 2008; 33(5):861–81. [PubMed: 18818425]
- Smith DB, Feng Z, Fennell ML, Zinn JS, Mor V. Separate and unequal: racial segregation and disparities in quality across U.S. nursing homes. *Health Aff (Millwood)*. 2007; 26(5):1448–58. [PubMed: 17848457]
- Stevenson DG. Nursing home consumer complaints and quality of care: A national view. *Medical Care Research and Review*. 2006; 63:347–68. [PubMed: 16651397]
- Welch WP, Miller ME, Welch HG, Fisher ES, Wennberg JE. Geographic variation in expenditures for physicians' services in the United States. *N Engl J Med*. 1993; 328(9):621–7. [PubMed: 8429854]

Appendices for Online Publication

Appendix A. Data

There are three primary sources of our data: the Minimum Data Set (MDS), Medicare Claims Records, and Online Survey of Claims (CASPER). We discuss each in turn.

MDS

MDS assessment forms are completed for all residents in certified nursing homes. MDS assessments are required for residents upon admission to the nursing facility and then periodically, within specific guidelines and time frames. MDS information is transmitted electronically by nursing homes to the MDS database in each state. MDS information from the state databases is captured into the national MDS database at Centers for Medicare and Medicaid Services (CMS).

Medicare Claims

For each patient, there is one record in a given year in the Medicare enrollment data set. Medicare is the federally funded program that provides health insurance for the elderly, persons with end-stage renal disease, and some disabled persons. For persons age 65 and over, 97 percent are eligible for Medicare. Almost all Medicare beneficiaries have Part A coverage that includes hospital, skilled-nursing facility, hospice, and some home health care. 96 percent of elderly Part A beneficiaries choose to pay a monthly premium to enroll in Part B of Medicare, which covers physician and outpatient services. While some Medicare

beneficiaries are enrolled in HMOs, most have fee-for-service (FFS) coverage. Information about Medicare eligibility and enrollment is available for all Medicare beneficiaries. Medicare claims (bills) are available only for persons with FFS coverage. For more information about Medicare data, please see the Centers for Medicare and Medicaid Services web site.

CASPER

Formerly known as OSCAR, CASPAR is a data network maintained by the Centers for Medicare and Medicaid Services (CMS) in cooperation with the state long-term care surveying agencies. CASPER is a compilation of all the data elements collected by surveyors during the inspection survey conducted at nursing facilities for the purpose of certification for participation in the Medicare and Medicaid programs. The CASPER database includes each nursing home's operational characteristics and aggregate patient characteristics for each facility. Onsite evaluations are conducted by state survey agencies at least once during a 15-month period, or as a result of a complaint being investigated.

Appendix B: Choice Sets

The first choice set is composed of any destination nursing home for any patient from a given zip code, which we call CS_{it}^1 , and is applied to all individuals in that zip code. We construct this set using the universe of patients who entered nursing homes in the year 2000. There are three potential problems with this choice set. First, the size of the choice set is limited by the number of patients observed. For example, if a zip code has just one patient entering a nursing home in 2000, then even if there are two nursing homes nearby, there will be only one nursing home in the choice set of that zip code. Second, this choice set is composed of revealed choices by the patients in a particular zip code. So if patients from a given zip code systematically avoid nearby poor quality nursing homes, one might underestimate the effects of distance for these individuals because the proximate nursing homes would not appear in the choice set. Finally, if neighborhoods of residence are already segregated by race or poverty, and individuals prefer to go to a nursing home with a higher proportion of their own race, then proximate nursing homes with a high fraction of other races might be excluded from the choice set. For example, if there are three nursing homes nearby and two are primarily white and the other is primarily black, then a proximate all-white zip code might have only the white nursing homes in their choice set, while the other nearby zip codes might have all three. In this case, it might not be evident that whites have a strong preference for racial homogeneity because the choice set for individuals in the white zip code would exclude the black nursing home.

Some of the limitations with the first choice set can be addressed with a distance-based choice set. A potential concern here, however, is that a distance that reflects a reasonably sized choice set in a high density area might be too small to capture a reasonable set of choices in a less dense area. Conversely, an appropriate distance in a less dense area may create an unrealistically large choice set in a high density area. Thus, the second choice set that we consider, CS_{it}^2 , is the union of the nursing homes within a 22-kilometer radius from the origin zip code and the geographically nearest 15 nursing home from the origin zip code.

The figure 22 km represents the 75th percentile in terms of distance travelled in our data. A disadvantage with this distance-based approach is that individuals who travelled more than that in our set criterion (i.e., they are in the upper 25% of the distance distribution and travelled to a nursing home that is the 16th in terms of distance from the origin zip code) must be omitted from the analysis.

Our third choice set, CS_{it}^3 , is the union of CS_{it}^1 and CS_{it}^2 . As this set is composed of both the sets, the estimated distance parameter with this choice set should not be contaminated by the distance criterion. We also impose a practical restriction that will permit us to account for state-level spatial correlation in a simple way—we remove from the choice set nursing homes from other states. In our sample, roughly five percent of all the patients went to an out-of-state nursing home. There are two possible reasons for going to a nursing home in another state. First, the patient may have moved to another state, say, to be close to family, and became a resident there, in which case, our measures of distance based on the original residence may be inappropriate. Second, the patient may have remained a resident of one state but ended up choosing a nursing home across a state border because it is a better match. In our data set, among all the individuals who went to a nursing home in a different state, half changed their residential zip code between 1998 and 2000 based on the Medicare address data. The comparable figure of the patients who went to a nursing home in the same state is roughly eight percent. These results at least suggest that a formal change in state residence underlies most cross-state movement and as such should be eliminated from the choice set for a given zip code.

Appendix Table 1

First stage: OLS regression of % of black among NH residents onto other NH characteristics

VARIABLES	(1) % of black in NH	(2) % of black in NH
% of black in the zip code of the NH	0.586*** [65.56]	0.556*** [64.30]
% of black in the nearest four zip codes of the NH	0.470*** [48.99]	0.469*** [50.92]
Number of CNAs per bed	-3.749*** [-7.049]	-4.105*** [-7.734]
Number of RNs per bed	-9.086*** [-8.425]	8.979*** [6.875]
Number of LPNs per bed	7.423*** [6.031]	10.22*** [8.429]
Average RUGS-III case-mix index	3.560*** [4.883]	3.432*** [4.617]
Occupancy rate	-3.373*** [-5.465]	-3.839*** [-6.338]
Total number of beds	0.0194*** [14.49]	0.0127*** [9.628]
% of Medicaid residents		0.112*** [26.48]
Multi-facility		0.765*** [4.526]
For profit		2.342*** [11.99]
Hospital based		0.749** [2.157]
ProPac acuity index		0.00151 [0.437]
Deficiency score		0.183*** [10.42]
Constant	0.588 [0.705]	-9.634*** [-11.22]
Observations	15,805	15,805
R-squared	0.678	0.703

Note: Robust t-statistics are in the brackets.

Appendix Table 2

Estimation results with alternative sample

	(1) Medicaid non-eligible	(2) Medicaid eligible	(3) Hip Fracture	(4) Heart failure	(5) Stroke
Distance from residential zip code	-0.084*** [-17.25]	-0.066*** [-16.30]	-0.083*** [-19.68]	-0.073*** [-12.94]	-0.075*** [-14.48]
Distance from originating hospital	-0.086*** [-18.26]	-0.063*** [-15.53]	-0.080*** [-18.52]	-0.079*** [-16.15]	-0.072*** [-14.34]
Number of CNAs per bed	-1.033*** [-7.996]	-1.409*** [-6.691]	-0.860*** [-5.397]	-1.147*** [-6.712]	-1.168*** [-7.280]
Number of RNs per bed	3.076*** [15.95]	2.085*** [8.179]	2.823*** [12.14]	2.431*** [9.852]	2.069*** [8.295]
Number of LPNs per bed	1.989*** [7.107]	1.285*** [5.001]	1.768*** [5.856]	2.060*** [5.577]	2.051*** [4.871]
% of black in NH	-0.045*** [-13.88]	-0.029*** [-11.53]	-0.049*** [-12.79]	-0.040*** [-14.37]	-0.043*** [-14.04]
Black * Distance from residential zip code	-0.00042 [-0.0507]	-0.00711 [-0.918]	-0.0036 [-0.248]	-0.0121 [-1.130]	0.00359 [0.386]
Black * Distance from originating hospital	-0.0212** [-2.332]	-0.0158* [-1.705]	-0.00955 [-0.599]	-0.0233** [-2.022]	-0.0071 [-0.758]
Black * Number of CNAs per bed	-0.108 [-0.450]	-0.109 [-0.408]	0.0827 [0.292]	-0.691 [-0.895]	0.284 [0.736]
Black * Number of RNs per bed	0.338 [0.946]	0.661** [2.340]	0.145 [0.257]	1.489** [2.536]	0.219 [0.414]
Black * Number of LPNs per bed	0.00429 [0.00864]	-0.0664 [-0.133]	0.261 [0.294]	-0.791 [-0.921]	0.137 [0.178]
Black * % of black in NH	0.0467*** [12.46]	0.0398*** [16.67]	0.0541*** [11.57]	0.0473*** [11.38]	0.0488*** [12.27]
Observations	3,451,269	914,088	409,164	179,381	254,765
N	93,200	25,112	11,801	4,814	7,179
% of black in sample	4.77	19.42	3.22	9.04	12.51

Note: All the regressions have total number of beds, occupancy rate, case mix index, and their interaction with black as additional independent variable.

Robust z-scores are in the brackets.

Appendix Table 3

Estimation results using individuals with different likelihood of becoming long-stay nursing home resident

Likelihood of becoming long-stay resident	(1) Quartile 1	(2) Quartile 2	(3) Quartile 3	(4) Quartile 4
Distance from residential zip code	−0.0914*** [−16.52]	−0.0830*** [−18.96]	−0.0830*** [−17.43]	−0.0759*** [−16.10]
Distance from originating hospital	−0.0978*** [−16.34]	−0.0827*** [−18.37]	−0.0785*** [−18.26]	−0.0710*** [−14.47]
Number of CNAs per bed	−1.350*** [−6.976]	−1.009*** [−6.316]	−1.028*** [−8.111]	−0.957*** [−5.689]
Number of RNs per bed	4.342*** [17.37]	3.026*** [15.50]	2.411*** [10.63]	1.447*** [5.297]
Number of LPNs per bed	1.894*** [5.353]	2.108*** [6.445]	2.083*** [7.533]	1.758*** [8.161]
% of black in NH	−0.0422*** [−10.79]	−0.0453*** [−12.25]	−0.0450*** [−13.02]	−0.0399*** [−14.81]
Black * Distance from residential zip code	0.0109 [1.168]	−0.00611 [−0.627]	0.00458 [0.446]	0.00153 [0.193]
Black * Distance from originating hospital	−0.0164 [−1.134]	−0.0231*** [−2.742]	−0.0180* [−1.690]	−0.00365 [−0.400]
Black * Number of CNAs per bed	−0.2 [−0.637]	−0.157 [−0.670]	0.105 [0.540]	−0.348 [−1.305]
Black * Number of RNs per bed	−0.00445 [−0.0140]	0.139 [0.269]	0.408 [0.924]	1.045*** [3.020]
Black * Number of LPNs per bed	−0.566 [−0.963]	0.192 [0.426]	−0.171 [−0.290]	−0.326 [−0.597]
Black * % of black in NH	0.0419*** [8.151]	0.0485*** [11.25]	0.0510*** [15.16]	0.0484*** [17.76]
Observations	1,073,370	1,035,213	1,052,818	1,095,612
N	28,268	28,645	28,805	29,124
% of black in sample	6.53	5.77	6.44	12.39

Note: Long-stay is defined as staying at a nursing home for more than 20 days following admission. The likelihood of becoming a long-stay resident has been estimated using a linear probability model and explanatory variables listed in Table 1 and primary diagnosis ICD9 fixed effects.

All the regressions have total number of beds, occupancy rate, case mix index, and their interaction with black as additional independent variable.

Robust z-scores are in the brackets.

Appendix Table 4

Estimation results using individuals from different hospitals

	(1) From 50% hospitals with low rehospitalization rate	(2) From 50% hospitals with high rehospitalization rate	(3) From hospitals with low share of black patients	(4) From hospitals with high share of black patients
Distance from residential zip code	-0.0826*** [-18.66]	-0.0788*** [-14.88]	-0.0886*** [-21.80]	-0.0777*** [-9.679]
Distance from originating hospital	-0.0845*** [-14.40]	-0.0754*** [-16.22]	-0.0890*** [-19.51]	-0.0745*** [-11.43]
Number of CNAs per bed	-1.108*** [-5.833]	-1.040*** [-7.991]	-1.089*** [-9.027]	-1.143*** [-5.172]
Number of RNs per bed	3.658*** [15.84]	2.282*** [11.28]	3.008*** [16.05]	2.576*** [9.765]
Number of LPNs per bed	1.796*** [5.192]	2.066*** [5.971]	2.056*** [6.615]	1.986*** [4.787]
% of black in NH	-0.0491*** [-8.622]	-0.0390*** [-12.98]	-0.0833*** [-15.11]	-0.0290*** [-10.58]
Black * Distance from residential zip code	0.000755 [0.0722]	0.00301 [0.472]	0.000621 [0.0766]	0.00238 [0.442]
Black * Distance from originating hospital	-0.0131 [-0.943]	-0.01 [-1.251]	-0.0218*** [-2.740]	-0.00992 [-1.381]
Black * Number of CNAs per bed	-0.613 [-1.578]	-0.103 [-0.385]	-0.0419 [-0.146]	-0.152 [-0.700]
Black * Number of RNs per bed	0.394 [1.098]	0.518 [1.451]	0.0296 [0.103]	0.587** [2.036]
Black * Number of LPNs per bed	-0.171 [-0.400]	-0.399 [-0.679]	-0.317 [-0.529]	-0.421 [-0.860]
Black * % of black in NH	0.0551*** [10.26]	0.0453*** [14.00]	0.0813*** [11.37]	0.0358*** [14.19]
Observations	1,797,450	2,567,907	2,290,414	1,763,067
N	54,374	63,938	62,103	39,645
% of black in sample	4.4	10.8	2.37	19.8

Appendix Table 5

Estimation results using individuals from different hospitals

	(1) South non-Metro region	(2) South Metro region	(3) Non-South Metro region
Distance from residential zip code	−0.0608*** [−21.42]	−0.0675*** [−9.054]	−0.0957*** [−16.80]
Distance from originating hospital	−0.0556*** [−20.78]	−0.0781*** [−10.67]	−0.105*** [−17.75]
Number of CNAs per bed	−0.614** [−2.457]	−1.904*** [−10.19]	−0.782*** [−5.893]
Number of RNs per bed	3.396*** [5.110]	2.676*** [12.51]	2.916*** [11.44]
Number of LPNs per bed	1.102 [1.391]	2.824*** [6.533]	1.899*** [5.531]
% of black in NH	−0.0224*** [−4.842]	−0.0381*** [−9.228]	−0.0521*** [−12.95]
Black * Distance from residential zip code	0.0104*** [3.108]	3.26E−05 [0.00679]	0.0209 [1.086]
Black * Distance from originating hospital	0.00980*** [3.257]	−0.00075 [−0.120]	−0.0614*** [−5.987]
Black * Number of CNAs per bed	−0.961* [−1.899]	−0.0268 [−0.0618]	−0.173 [−0.866]
Black * Number of RNs per bed	1.184 [1.582]	0.189 [0.332]	0.163 [0.514]
Black * Number of LPNs per bed	0.512 [0.469]	−1.962*** [−3.222]	0.591 [1.157]
Black * % of black in NH	0.0412*** [5.796]	0.0429*** [9.778]	0.0535*** [10.84]
Observations	1,797,450	2,567,907	2,290,414
N	10,035	28,110	62,367
% of black in sample	11.74	12.07	6.56

Note: There are very few black patients in non-south, non-metro areas, and the estimation did not converge.

Appendix Figure 1.

Change in likelihood of being admitted if % of black in nursing home increases by 10 points

Appendix Figure 2.

Location of the intervened nursing homes for second simulation

Highlights

- Patients of each race prefer nursing homes with higher prevalence of patients of their own race.
- Preference for distance and quality of care are same for patients of both the races.
- Both preference for distance and preference for racial homogeneity contribute to racial disparity in nursing home quality of care.
- Results are robust among subgroup of patients with same Medicaid eligibility and similar likelihood of becoming long-stay resident.
- Simulations based on estimated sorting model suggest that interventions targeting nursing homes serving minority population would yield a greater reduction in racial quality disparities than interventions targeting low quality nursing homes.

Figure 1.
Simulated NH % Black by NB % Black at NH Location and Preference Type

Figure 2.
Simulated Density of Racial Composition of NH for Person from Black Majority NB by
Race and Preference Type

Figure 3.
Simulated NH-NB Distance by Patient NB % Black. Race and Preference Type

Figure 4.

Simulated Average NH % Black by Patient NB % Black. Race and Preference Type

Figure 5.
Simulated NH Quality by Patient NB % Black, Race and Preference Type

Figure 6.
Polynomial regression of NH-NB distance onto log of % Blacks in NB zip code

Figure 7.

Polynomial regression of average NH % Black onto log of % Blacks in NB zip code

Figure 8.
Polynomial regression of adjusted number of RNs per bed by % of Black in NB zip code

Figure 9.
Spatial model of NHs and NBs

Figure 10.
Simulation of case where no individual has race preference

Figure 11.
Simulation with nursing staff based intervention

Figure 12.
Simulation with race-based intervention

Table 1

Descriptive statistics of individual characteristics

	White N= 596,419 (92.2%)		Black N= 50,075 (7.8%)	
	Mean	Std. Dev.	Mean	Std. Dev.
Individual characteristics				
Age	80.87	7.55	78.69	8.21
Male	0.351	0.477	0.363	0.481
Medicaid eligibility	0.183	0.387	0.512	0.500
High school graduate	0.334	0.472	0.265	0.441
More than high school	0.323	0.468	0.222	0.415
Missing education	0.155	0.362	0.143	0.350
Elixhauser comorbidity index	2.48	1.34	2.58	1.28
Deyo comorbidity index	1.59	1.86	1.89	1.99
Length of stay at hospital	9.11	7.44	11.40	9.96
Stayed for more than 20 days at NH	0.61	0.49	0.69	0.46
Residential zip code characteristics				
% of black	4.65	9.64	41.01	31.06
% of rural population	23.96	33.12	16.59	30.33
Per capita income	23523	9320	21642	9800
Poverty rate	8.63	5.16	12.98	8.08

Table 2

Descriptive statistics of treating nursing home characteristics

Variable	White		Black	
	Mean	Std. Dev.	Mean	Std. Dev.
Distance from residential zip code (in kilometers)	15.81	24.04	13.18	17.88
Distance from originating hospital (in kilometers)	12.30	25.54	9.38	19.31
Percent of residents that are black	6.66	10.97	35.22	27.34
Number of CNAs per bed	0.401	0.148	0.389	0.114
Number of RNs per bed	0.115	0.129	0.098	0.119
Number of LPNs per bed	0.152	0.090	0.159	0.077
Average RUGS-III case-mix index ¹⁸	1.08	0.08	1.08	0.08
Total number of beds	118.41	80.71	139.71	99.49
Occupancy rate	0.85	0.14	0.86	0.13

Note: CNA, Certified Nursing Assistant; RN, Registered nurse; LPN, licensed practical nurse; RUGS, Resource Utilization Group.

¹⁸Annual facility avg. RUG-III Normalized Case-Mix Index (NCMI) at admission: The Resource Utilization Groups (RUG-III) case-mix algorithm was developed to provide a patient-specific means of allocating health care resources based on the variable costs of caring for individuals with different needs. The current Version 5.12 of RUG-III uses 108 variables from the MDS to create 44 or alternatively 34 categories of patients with homogeneous resource use patterns. The RUG-III algorithm explains about 55% of variance in resource use, and it has been validated in a number of countries through a series of international studies.

Table 3

Comparison of the choice sets: MNL estimation of choice decision

VARIABLES	(1) CS_{it}^1	(2) CS_{it}^2	(3) CS_{it}^3	(4) CS_{it}^3 with same state NHs
Distance from residential zip code	−0.0599*** [−167.4]	−0.142*** [−225.3]	−0.0796*** [−207.2]	−0.0804*** [−17.90]
Distance from originating hospital	−0.0631*** [−193.2]	−0.0882*** [−175.5]	−0.0806*** [−234.3]	−0.0794*** [−17.45]
% of black in NH	−0.0285*** [−73.80]	−0.0390*** [−98.67]	−0.0376*** [−105.0]	−0.0377*** [−14.01]
Black * Distance from residential zip code	0.00686*** [4.681]	0.0112*** [4.842]	0.0017 [1.089]	0.0039 [0.561]
Black * Distance from originating hospital	−0.00554*** [−4.178]	−0.0138*** [−7.112]	−0.00938*** [−6.684]	−0.0085 [−0.975]
Black * % of black in NH	0.0366*** [56.41]	0.0454*** [69.30]	0.0449*** [75.22]	0.0449*** [15.46]
Pseudo R squared	0.2029	0.2773	0.2487	0.2419
Observations	2,214,883	3,816,369	4,843,817	4,367,918
# of individuals	120,566	120,566	120,566	120,566
# of individual dropped	531	11913	0	2256
# of observation dropped	531	341221	0	38904
Error clustered by state	N/A	N/A	N/A	Yes

Notes: All the regressions have total number of beds, occupancy rate, total number of CNAs, RNs, and LPNs, case mix index, and their interaction with black as an additional independent variable.

Table 4

Estimation of the choice function

	(1)	(2)	(3)	(4)	(5)
Distance from residential zip code	-0.0787*** [-16.60]		-0.0796*** [-16.97]	-0.0797*** [-16.90]	-0.0808*** [-17.00]
Distance from originating hospital	-0.0852*** [-17.70]		-0.0806*** [-17.54]	-0.0809*** [-17.48]	-0.0785*** [-16.90]
Number of CNAs per bed		-0.753*** [-7.001]	-1.032*** [-8.207]	-1.047*** [-8.398]	-0.585*** [-5.659]
Number of RNs per bed		2.791*** [15.92]	2.958*** [15.96]	2.921*** [15.62]	0.178 [1.112]
Number of LPNs per bed		1.535*** [6.501]	1.844*** [6.823]	1.881*** [6.885]	0.868*** [4.451]
% of black in NH		-0.0346*** [-12.35]	-0.0376*** [-14.92]	-0.0419*** [-14.52]	-0.0335*** [-11.74]
Residual				0.00780*** [3.853]	0.0136*** [6.357]
Black * Distance from residential zip code	0.000575 [0.0806]		0.0017 [0.236]	0.000992 [0.138]	0.000876 [0.120]
Black * Distance from originating hospital	-0.00913 [-0.998]		-0.00938 [-1.074]	-0.00982 [-1.096]	-0.01 [-1.114]
Black * Number of CNAs per bed		-0.878*** [-3.883]	-0.273 [-1.188]	-0.257 [-1.136]	-0.177 [-0.962]
Black * Number of RNs per bed		0.0318 [0.117]	0.148 [0.492]	0.189 [0.641]	0.477 [1.401]
Black * Number of LPNs per bed		0.701 [1.586]	-0.234 [-0.500]	-0.241 [-0.520]	-0.175 [-0.478]
Black * % of black in NH	0.0416*** [15.78]		0.0449*** [16.12]	0.0480*** [14.96]	0.0468*** [19.28]
Black * Residual				-0.00464* [-1.775]	-0.0098*** [-4.210]
Observations	4,367,918	4,367,918	4,367,918	4,365,357	4,365,357

Notes: In all, the regressions are based on 118,286 individuals with 7.88% black.

(2–4) have total number of beds, occupancy rate, case mix index, and their interaction with black as additional independent variable. The residuals come from regression presented in Column 1 of Table 5.

(5) includes total number of beds, occupancy rate, share of Medicaid patients, multi-facility dummy, profit status dummy, hospital based dummy, NCMI, and a severity index of the patients and their interaction with black as additional independent variable. Residuals come from regression presented Column 2 of Table 5.

Robust z-scores are in the brackets.

Table 5

Simulation results of segregation and quality gap with alternative race preference parameter: Regression of average destination NH attributes onto race dummy.

Dependent variable		(1) Predicted	(2) Predicted without race preference
% of black in the NH	Uncontrolled	28.57*** [305.3]	20.85*** [235.5]
	Controlling for NB zip code	14.29*** [160.3]	1.055*** [13.53]
Adjusted number of RNs per beds	Uncontrolled	−0.0113*** [−18.13]	−0.00413*** [−6.496]
	Controlling for NB zip code	−0.0217*** [−24.01]	−0.00682*** [−8.983]

Simulation results of segregation and quality gap under alternative interventions: Regression of average destination NH attributes onto race dummy.

Table 6

Dependent variable	Nursing staff based intervention		Race based intervention	
	(2) Before sorting	(3) After sorting	(4) Before sorting	(5) After sorting
Change in % of black in the NH	Uncontrolled	2.47*** [138.11]		2.41** [134.14]
	Controlling for NB zip code	1.96*** [112.83]		1.88*** [108.02]
Adjusted number of RNs per beds	Uncontrolled	-0.029*** [-2.20]	2.32*** [221.16]	1.9*** [22.01]
	Controlling for NB zip code	0.059*** [7.60]	1.23*** [191.05]	-0.31*** [-5.63]