Why Are Minorities Disproportionately Concentrated in Low-Quality Nursing Homes?

Alden Cheng

May 22, 2024

Preliminary: please do not circulate.

Abstract

This paper studies the underlying mechanisms giving rise to high levels of racial segregation and disparities across nursing homes in the US. Descriptively, I find that while residential segregation is an important explanation for racial segregation across nursing homes, it struggles to explain disparities, as defined by the difference in average quality of nursing homes that Whites and minorities go to. I then provide reduced form evidence for several other explanations for segregation and disparities: nursing homes seem to discriminate against minorities in their admission practices, individuals tend to choose nursing homes with a higher share of residents of their own race, and minorities seem less sensitive to nursing home quality. Next, to disentangle and quantify the effects of these explanations, I estimate a structural model and conduct counterfactual simulations. Estimates from the structural model provide support the reduced form evidence, while counterfactual simulations suggest that residential segregation is the single most important force behind nursing home segregation, although there seems to be complementarities between various explanations. The simulations also suggest that the importance of various explanations for disparities differs widely across states, and may even have opposite effects for different minority groups.

1 Introduction

Racial segregation is a pervasive phenomenon in a number of important settings, such as school

I am very grateful to Mackenzie Alston, Josh Angrist, David Autor, Zarek Brot-Goldberg, John Bowblis, David Card, Bryan Chu, Viola Corradini, Amy Finkelstein, Ashvin Gandhi, Jon Gruber, Youjin Hahn, Ben Handel, Jisoo Hwang, Simon Kosali, Riley League, Jungmin Lee, David Molitor, Kosali Simon, Hee-Seung Yang and seminar participants at University of Illinois Urbana-Champaign, Seoul National University, and Yonsei University for their valuable comments.

(Billings, Deming and Rockoff 2014), neighborhood (Card, Mas, and Rothstein 2007), and healthcare institutions (Baicker et al. 2004). Moreover, in many of these cases, minorities tend to be disproportionately concentrated in lower-quality institutions, leading to concerns over racial disparities.

However, despite the vast literature on segregation and disparities, there remains important unanswered questions. First, most rigorous studies tackle a single explanation in isolation, producing credible evidence on its existence. Yet, it is difficult to compare the relative importance of different channels across various studies and settings, and there may be potential complementarities between different explanations that cannot be identified by this approach.

Second, discussions around segregation and racial disparities do not always make a clear distinction between the two. To understand why they may differ, suppose we define disparities as the difference in average quality of hospitals that different races go to. Other than the extreme case of perfect integration where disparities are be zero, there is no mechanical relationship between the two.[1] If the importance of various explanations for segregation and disparities differ substantially, policymakers will need to think carefully about whether their main goal is to reduce segregation or disparities when deciding which policies to prioritize.

In this paper, I take a first step to filling these gaps in the research by studying segregation and racial disparities in the nursing home context, taking advantage of a rich administrative data set on the universe of nursing home residents. I start by establishing reduced form evidence on several potential causes of segregation and disparities in nursing homes, namely residential segregation, discrimination, in-group preferences, and heterogeneity in sensitivity to quality. To more cleanly disentangle these explanations and quantify their contributions, I then estimate a structural model and conduct counterfactual simulations.

First, while simple descriptive facts suggest that residential segregation is likely an important contributor to nursing home segregation, it is unlikely to fully explain racial disparities, as measured by the gap in average quality of nursing homes chosen by different racial groups. Most residents choose a nursing home relatively close to where they live, so residential segregation is positively correlated with nursing home segregation. Yet, even conditional on zip code of prior residence, minorities still tend to be admitted to lower-quality nursing homes.

Second, I find that nursing homes seem to discriminate against minority residents in their admissions

[1] In particular, even with complete segregation, it is still possible that disparities may be zero. To see why, suppose we randomly chose hospitals to completely populate with individuals of a single race. Then, disparities with be zero in expectation.

practices: they are less likely to admit minority residents during times of capacity strain (when the option value of an empty bed is higher).[2] If higher-quality nursing homes are in higher demand and experience greater capacity strain, such admissions practices may give rise to racial disparities.[3]

Third, residents seem to prefer nursing homes with a higher share of residents of their own race: dynamic panel and event study estimates both show that a shock to the share of minority admissions to a nursing home has a persistent effect, consistent with predictions from a tipping point model (Schelling 1971; Card, Mas, and Rothstein 2007). While a persistent supply-side change may also give rise to the same pattern, I find that short-run fluctuations in racial composition predicts the race of admissions in the near future as well. Such in-group preferences may have an effect not only on segregation, but may also perpetuate or even exacerbate racial disparities over time: given an initial distribution with minorities disproportionately living in low-quality nursing homes, minorities' demand for such nursing homes will be higher, leading to an even larger share living in these nursing homes, and so forth.

Fourth, I find suggest evidence that minorities may be less sensitive to nursing home quality than Whites are. An increase in quality at a given nursing home increases the probability that future residents will be White and reduces the probability that they will be Black, even after controlling for racial composition of said nursing home, local demographic trends, and the zip code of residents' prior address.

A shortcoming of the reduced form evidence is that disentangling different explanations is challenging. For example, observed choices may either reflect residents' preferences or rejections by other nursing homes (which are not observed in the data). Hence, to more cleanly disentangle these channels, in the second part of the paper I estimate a structural model for a few of the largest states in my data — California, Texas, Florida, and New York — between 2008 and 2010.[4] The structural model incorporates distance, racial preferences, and heterogeneous sensitivity to quality in residents' decision utility, as well as occupancy, race, and other resident characteristics in nursing homes' admissions rule. Estimates from the model support the explanations considered in the reduced form evidence: discrimination, in-group preferences, and information frictions all seem to play some role.

[2] I do not take a stance on whether this behavior is due to taste-based (Becker [1957] 1971) or statistical discrimination (Arrow 1972a, b) of a profit-maximizing firm (although I do control for other resident characteristics that may be associated with both race and their profitability to nursing homes).

[3] It is important to emphasize that I only focus on one specific form of discrimination. For example, information frictions that minority residents face may be the result of systematic discrimination more broadly (e.g. education, and internet access). Similarly, while I measure disparities based on nursing home *choice*, provision of lower-quality care to minorities and selective discharge practices by nursing homes (Hackmann, Pohl, and Ziebarth 2020) may further contribute racial disparities in resident *outcomes*.

[4] I estimate the structural model separately for each state due to computational constraints. The four states were chosen because they had the largest sample size among states with a decent number of Black and Hispanic residents in my sample.

A second advantage of the structural model is that it allows me to quantify the contributions of different explanations to segregation and disparities via counterfactual simulations. These simulations indicate that residential segregation is the single most important explanation for statewide segregation, a result consistent across different states. However, there seems to be complementarities between residential segregation and other explanations for segregation, such as in-group preferences. In particular, while eliminating discrimination or in-group preferences by themselves have very little effect on segregation, they become much more effective in the absence of residential segregation.

In addition, the counterfactual simulations show that the importance of different explanations for disparities vary widely across states. Moreover, the simulations suggest that eliminating residential segregation may even hurt minorities in some cases: for example, it lowers the average quality of nursing homes that Hispanics are admitted to in Florida even as it reduces Black-White disparities, which also highlights the potential for policies aimed at reducing disparities to have opposite effects on different minority groups.

This paper is linked to a vast literature on racial segregation and disparities. A number of previous studies have produced credible evidence on various explanations for these patterns in a number of settings. For example, Card, Mas and Rothstein (2007) show support for one of the key implications of Schelling's model of in-group preferences — the existence of tipping points — in the context of neighborhood choice, while Aaronson, Hartley, and Mazumder (2021) and Derenoncourt (2022) highlight the role of discrimination and location in the evolution of segregation and disparities. In addition, Billings, Deming, and Rockoff (2014) establish a direct link between segregation and outcomes disparities using a natural experiment induced by the end of race-based busing in Charlotte-Mecklenburg schools (CMS), and Oosterbeek, Sóvágó, and van der Klaauw (2021) document how heterogeneous "preferences" may interact with school segregation in Amsterdam.

My paper contributes to this literature by studying several of these explanations simultaneously within a single setting, since most previous studies have focused on one explanation in isolation. This allows me to compare their effects on segregation and disparities, as well as potential interaction effects between various counterfactual policies. Moreover, by studying multiple forces in concert, I show that policies which reduce segregation may be different than those that reduce racial disparities, and may even have effects of opposite signs on segregation and disparities (as well as for different minority groups when it comes to disparities).[5,6]

More narrowly, this paper is related to a literature on racial segregation and disparities in healthcare settings (Baicker et al. 2004; Smith et al. 2007; Rahman and Foster 2015). With the exception of the last reference, most of these studies tend to be reduced form and descriptive. As I discuss in greater detail later in the paper, reduced form evidence often struggles to disentangle demand and supply side explanations, so in this study I estimate a structural model in order to identify the underlying mechanisms. The closest study to this paper is Rahman and Foster, who also study the role of in-group preferences, location, and preference heterogeneity in the context of nursing homes. However, they do not model nursing homes' admissions decisions, so their structural estimates may reflect both residents' preferences and nursing homes' admission decisions.

This paper proceeds as follows. In section 2, I provide some background on nursing homes and introduce my main data sources, before laying out some descriptive statistics on racial segregation and disparities. In section 3, I present reduced form evidence for various explanations for segregation and disparities. In section 4, I introduce and estimate an empirical matching model that incorporates all of these elements, and in section 5, I conduct counterfactual simulations to quantify the importance of these factors for explaining racial segregation and disparities. Section 6 concludes.

2 Background

There are roughly 15,000 nursing homes in the US providing care for about 1.3 million Americans (CDC), and an estimated 56 percent of Americans aged 57–61 are expected to spend at least one night in a nursing home during their lifetimes (Hurd, Michaud, and Rohwedder 2017). However, a large literature has documented substantial segregation across nursing homes (see the meta-analysis by Mack et al. 2020), and that minorities tend to disproportionately choose lower-quality nursing homes (Li et al. 2015). Quality of nursing homes vary widely and can have meaningful impacts on

[5] Clearly, under perfect integration, the average quality of nursing homes chosen by majority and minority groups will be the same. However, a reduction in segregation that does not achieve full integration need not necessarily reduce the gap in quality chosen by majority and minority groups. To see how this may be possible, consider a simplified example with 3 nursing homes, A, B, and C, with cardinal quality measures given by 5, 2, and 1 respectively. There are 200 residents from the majority group, and 100 residents from the minority group, and initially all minority group residents reside in nursing home B, whereas majority group residents are split equally between nursing homes A and C. Now, suppose that 50 minority residents move from nursing home B to C, and 50 majority residents move from nursing homes C to B. This achieves a reduction in segregation, but the average gap in the quality of nursing homes chosen by majority and minority groups increases from $(5+1)/2-2=1$ to $(5/2+2/4+1/4)-(2+1)/2=7/4$.

[6] It should be noted that there are some crucial differences between the nursing home setting and those that were studied previously, such as education. Perhaps most meaningfully, peer effects are less likely to play an important role in nursing home settings compared to education: the racial composition of residents in an individual's nursing home is unlikely to have a direct causal impact on her own health. Such differences should be kept in mind when generalizing the findings in this paper to other settings.

important outcome such as short-term mortality (Cheng 2023), so these patterns of racial segregation and disparities paint a worrying picture when it comes to tackling racial gaps in healthcare. Finally, most residents are covered (at least in part) by insurance (most often Medicare or Medicaid), so differential distance to nursing homes is typically a much more important factor in residents' nursing home choice compared to differences in out-of-pocket prices.

2.1 Data

The primary data source for this paper is the Minimum Data Set 2.0 (MDS). All nursing homes that receive federal funding are required to fill out MDS assessment forms at regular intervals (42 CFR §483.20).[7,8] Data collected from the MDS assessments includes information on residents' demographics, cognitive status, communication and hearing patterns, vision patterns, mood and behavior patterns, psychosocial well-being, physical functioning and structural problems, continence issues, disease diagnoses (including ICD-9 codes), health conditions, oral health, nutrition, dental status, skin conditions, activity pursuit patterns, medications, special treatments and procedures, and discharge potential.

I supplement the MDS with data on nursing homes from other sources. This includes the Online Survey Certification and Reporting (OSCAR) surveys (which contain information such as nursing homes' ownership status and staffing levels), data on deficiency citations, and five-star ratings for nursing homes.[9] For my main analysis, I consider a composite quality index, based on RN (registered nurse) staffing, LPN (licensed practical nurse) staffing, and (the negative of) standard deficiencies (which are deficiency citations given by inspectors either during their annual visits). I standardize each of these components (so they all have mean zero and unit variance), before taking their sum and standardizing the sum to create the final quality index. There are several other potential quality measures that I do not use for various reasons. Probably the most well-known nursing home quality

[7] The set of nursing homes receiving federal funding account for roughly 96 percent of all nursing homes (Grabowski, Gruber, and Angelelli 2008).

[8] Assessment forms must completed upon admission, at discharge (or death), quarterly in between, and whenever there is a significant change in status. MDS forms are typically filled out by a registered nurse (RN), or at least certified by one. Any willful misrepresentation in the MDS forms may result in penalties under the False Claims Act. This is not limited to upcoding and variables that affect reimbursements directly but also other variables related to resident well-being. This is because nursing homes "must provide services to attain or maintain the highest practicable physical, mental, and psychosocial well-being of each resident" (42 CFR §1395i–3) to be certified to receive federal funding. Hence, any misrepresentation pertaining to resident wellbeing may be interpreted as being related to misrepresentation connected to a requirement for federal funding, and thus falls under the False Claims Act. Moreover, several studies on the accuracy of MDS data have found it to be fairly reliable (Shin and Scherer, 2009).

[9] The OSCAR data is available from 2000 onwards from LTCFocus.org, which is maintained by Brown University Center of Gerontology and Healthcare Research. LTCFocus is sponsored by the National Institute on Aging (1P01AG027296) through a cooperative agreement with the Brown University School of Public Health. Data on deficiencies, Medicare cost reports, and five-star ratings are available from the CMS website.

measure is the star rating provided by the CMS, but it is only available from the end of 2008 onwards, whereas my sample ranges from 2000–2010. Similarly, in addition to annual inspections, inspectors may also visit a nursing home in response to a complaint, and I term deficiency citations from such visits as "complaint deficiencies". However, this measure is only available from 2006 onwards. Nonetheless, I check that my main results are robust to individual quality measures, leaving most of these results in the appendix. For more details on the data used in this paper, see Appendix section B.

2.2 Sample and Summary Statistics

When possible, I use the entire sample of residents in the US. However, this is computationally infeasible for some of the analysis, particularly for the structural model. In these cases, I focus on residents in the four of the largest states in my sample — California, Texas, Florida, and New York — from 2008–2010.

Table 1 shows summary statistics for nursing homes residents in the entire US, overall and separately by race. We observe that residents are typically White, female, advanced in age, and have less than a bachelor's degree. Moreover, most of them are admitted from acute care hospitals (and are thus likely to be short stay), 8 percent die within 90 days of admission, and 22 percent are already diagnosed with dementia at admission. Comparing characteristics of residents from different racial groups, the most notable differences are that White residents are on average older, more educated, and more likely to die within 90 days of admission, compared to Black and Hispanic residents. This last fact may seem somewhat surprising, but may be explained by minority residents being substantially younger (and hence, likely in better health) than White residents at admission.[10] Appendix Table A.1 shows these summary statistics by state for the four states used in the structural estimation.

Next, Table 2 shows summary statistics for nursing homes across the US in column 1, weighted by the number of residents admitted between 2008–2010). Nursing homes have over a hundred beds on average, and occupancy rates are often above 80 percent. In addition, more than half of nursing homes are owned by chains, and more than 60 percent are for-profit. I also show summary statistics for various quality measures, including staffing hours for RNs, LPNs and Certified Nursing Assistants (CNAs), and standard deficiencies.[11]

[10] One potential reason for why White residents are older at admission may be that they have better resources which allow them to avoid going to nursing homes unless truly necessary.

[11] Specifically, in Table 2 I show standard deficiency citations. since data on complaint deficiencies are only available from 2006 onwards.

2.3 Broad Patterns of Racial Segregation and Disparities

To measure racial segregation across nursing homes, I use the index of dissimilarity. This index measures segregation across two racial groups and lies between 0 and 1, representing perfect integration and complete segregation respectively. For two groups A, and B, the dissimilarity index D of some geographical region is defined by:

$$D = \frac{1}{2}\sum_{j}\left|\frac{a_j}{\sum_{j'} a_{j'}} - \frac{b_j}{\sum_{j'} b_{j'}}\right|,$$

where a_j (respectively b_j) is the number of residents of group A (B) in nursing home j. An interpretation of the index D is that it is the proportion of one of the two groups that would have to move to different nursing homes in order for the distribution of the groups in each nursing home to match the overall distribution of these groups in the geographical region under consideration. Since the index of dissimilarity is only defined for two distinct racial groups, I compute this measure separately based on Black versus non-Black racial groups, and Hispanic versus non-Hispanic racial groups. Focusing on the distribution of state-level nursing home segregation shown in Figure 1, we observe that dissimilarity indices for most states range from 0.3 to 0.7, consistent with previous research finding that nursing home segregation are similar to those for residential segregation (Mack et al. 2020).

To measure racial disparities across nursing homes, I regress quality of the nursing home that residents are admitted to on race dummies, taking the coefficient estimates on the race dummies β^q_{black} and $\beta^q_{hispanic}$ as the racial gaps:

$$q_{j(i)} = \beta^q_0 + \beta^q_{black} black_i + \beta^q_{hispanic} hispanic_i + \epsilon^q_i, \tag{1}$$

clustering standard errors at the nursing home level. As my main quality measure, I use a standardized quality index based on registered nurse (RN) staffing, licensed practitioner nurse (LPN) staffing, and fewer standard deficiencies (Kling, Liebman, and Katz 2007). Figure 2 shows the racial gap differs substantially across states, but a general pattern emerges whereby Black residents tend to stay in lower-quality nursing homes, whereas the Hispanic-White gap is much smaller.[12] Appendix Figures A.2, A.4, A.5, and A.6 show that the same patterns hold qualitatively when we use RN staffing, LPN staffing, fewer standard deficiencies, fewer complaint deficiencies, or star ratings individually as the quality measure.

[12] Coefficient estimates with large standard errors due to small sample sizes are omitted for legibility purposes.

However, despite substantial statewide racial segregation and disparities, the cross-sectional relationship between the two are weak. Indeed, the scatterplots in Figure 3 shows that there is at best a weak cross-sectional relationship between the estimated racial disparities and segregation at the state level, a possibility explained in the introduction. This absence of a clear cross-sectional relationship persists when we consider racial disparities based on other measures, such as RN staffing, LPN staffing, fewer standard deficiencies, and fewer complaint deficiencies, as Appendix Figures A.7, A.8, A.9, and A.10 show.

3 Reduced Form Evidence

In this section, I present several pieces of reduced form evidence on the potential causes of segregation and disparities across US nursing homes.

3.1 Residential Segregation is Linked to Nursing Home Segregation, but is Unlikely to Explain Disparities

If neighborhoods are highly segregated and residents prefer not to travel long distances to nursing homes, then neighborhood segregation feeds into nursing home segregation. As a first piece of evidence that residential segregation feeds into in nursing home segregation, in Figure 5 I show a binscatter of MSA-level Black-White dissimilarity indices for residential segregation against state-level Black/non-Black nursing home dissimilarity indices,[13] and we observe that there is indeed a strong positive correlation between residential and nursing home segregation.

Next, I show that after accounting for residential segregation, measured nursing home segregation becomes much smaller. In particular, given that most residents choose a nursing home relatively close to them, I compute dissimilarity indices based on a 15-mile radius of each 5-digit zip code for any resident's prior residential address and compare this to the state-level dissimilarity indices. To understand what this comparison tells us, consider two extremes as a thought experiment. If distance to nursing homes does not matter for residents choosing their nursing homes and residents are willing to travel to any nursing home within their own state, then this local measure of segregation will be identical to the statewide measure. By contrast, if distance does matter to residents, and residential segregation is the only source of residential segregation, then the local segregation measure will be

[13] Data on MSA-level Black-White dissimilarity indices was downloaded from https://censusscope.org/us/rank_dissimilarity_White_Black.html on October 15, 2023.

close to zero. In fact, Figure 4 shows that the truth lies somewhere in between these two extremes: dissimilarity indices are almost all smaller than 0.4, as compared to 0.3–0.7 at the state level, which suggests that residential segregation is an important explanation for overall nursing home segregation, but dissimilarity indices for many zip codes are also significantly different from zero, rejecting the notion that residential segregation is the sole cause.

To show the extent to which residential segregation can explain disparities, I estimate equation (1) with fixed effects for the zip code where residents used to live before being admitted to a nursing home. The results in Table 3 show that conditioning on zip code of prior residence reduces the racial gaps (relative to the unconditional gaps), but minority residents still tend to be admitted to lower-quality nursing homes, so residential segregation is unlikely to entirely explain racial disparities. Appendix Table A.2 shows that the same qualitative patterns hold when considering individual quality measures, such as staffing, fewer standard and complaint deficiencies, and nursing homes' 2009 star ratings.

3.2 Discrimination by Nursing Homes May Give Rise to Disparities

Geographical proximity is not the only potential barrier to access: nursing homes may also discriminate against minorities in their admissions process. In fact, Gandhi (2020) shows that when capacity is strained, nursing homes tend to become more selective in the types of residents they admit and are less likely to admit Medicaid residents (who tend to be less profitable), presumably because the option value of an empty bed is increasing in capacity strain. Nursing homes may also find minority residents less attractive if certain minority characteristics are negatively correlated with profitability (e.g., Medicaid status) or due to outright taste-based discrimination. If this is the case and higher-quality nursing homes experience greater demand, then these selective admissions practices may give rise to the observed racial disparities.[14]

To probe this possibility, I test two predictions from Gandhi's model. First, due to capacity constraints, nursing homes should admit fewer new residents when they are close to capacity. Second, and more importantly, the characteristics of residents that nursing homes admit during times of high and low occupancy should differ, given that nursing homes are more selective when they are closer to capacity.

[14]This is related to findings on discrimination in the rental market, where Christensen and Timmons (2023) find that the gap in response rates to minority and White individuals in a correspondence study is greater when demand for the rental property is higher.

To test the first prediction, I run regressions at the nursing home-day level of the form:

$$admissions_{jd} = \alpha^{cap} + \beta^{cap} occupancy_{jd} + \delta^{cap}_{jm} + \epsilon^{cap}_{jd},$$

where $admissions_{jd}$ and $occupancy_{jd}$ are respectively measures of new admissions, and lagged seven-day average of some occupancy measure for nursing home j on day d, controlling for nursing-home month fixed effects δ^{cap}_{jm} in order to isolate temporary occupancy fluctuations (as opposed to longer term expansions and contractions in capacity). The results in Table 4a indicate that nursing homes indeed admit fewer residents when occupancy is higher than usual, and Appendix Table A.4 shows that these results are robust to different measures of new admissions.

I test the second prediction by running regressions at the resident level of the form:

$$x_{ip} = \alpha^{select}_p + \beta^{select}_p occupancy_{j(i),d(i)} + \gamma^{select}_{\sim p} x_{i,\sim p} + \delta^{select}_{j,p} + \epsilon^{select}_{ip},$$

where x_{ip} is some characteristic p of resident i, $occupancy_{j(i),d(i)}$ is a measure of the nursing home when it admitted i, controlling for nursing home fixed effects $\delta^{select}_{j,p}$, and either controlling for other resident characteristics $x_{i,\sim p}$ or not in different specifications.

The results in Tables 4b and 4c indicate that when occupancy is higher than usual, nursing homes are more likely to admit post-acute care residents, given that they are often covered by Medicare, which has higher reimbursement rates than Medicaid. Moreover, they are also less likely to admit Black and Hispanic residents, a pattern persists even after controlling for other resident characteristics, which stronger results for Black residents. Appendix Table A.5 shows that the same qualitative patterns generally hold using other measures of nursing home occupancy.

3.3 In-Group Preferences May Explain Segregation and the Perpetuation of Disparities

Another potential explanation for why minorities continue choosing lower-quality nursing homes despite the presence of higher-quality nursing homes nearby is in-group preferences (Schelling 1971). For example, individuals from a given racial group may prefer to interact with other members of the same racial group due to shared experiences, or because they believe they will be treated with more respect. If this is the case, then an initial distribution of minorities concentrated in lower-quality nursing homes may persist moving forward even in the absence of other inequities.

A key prediction of models of in-group preferences is that a shock to the minority share may lead to an equilibrium switch: specifically, a positive shock to the share of a racial group r at a nursing home may have persistent effects, as future residents of race r find that nursing home more attractive while residents of other races r' find it less attractive (Card, Mas and Rothstein 2007; Billings, Deming and Rockoff 2014; Hailey 2022). While the sample sizes for individual nursing homes are too small to conduct a tipping point analysis of Card, Mas and Rothstein (who do so in the neighborhood context using Census tract data),[15] here I present evidence using dynamic panel methods and event study-type analyses.

I start by estimating autoregressive models at the nursing home-year level, based on the share of admitted residents who are of a minority group (either Black or Hispanic):

$$s_{jt}^{r} = \alpha^{r,ingroup} + \beta^{r,ingroup} s_{j,t-1} + \delta_{j}^{r,ingroup} + \gamma_{ct}^{r,ingroup} + \epsilon_{jt}^{r,ingroup}, \qquad (2)$$

where s_{jt}^{r} is the share of residents admitted to nursing home j that are of race r, and observations are weighted by the number of admissions the nursing home j receives in year t. I control for time-invariant factors affecting share of Black admissions by including nursing home fixed effects $\delta_{j}^{r,ingroup}$, and demographic trends by including county-year fixed effects $\gamma_{ct}^{r,ingroup}$.

The OLS estimates of equation (2) in column 1 of panels A and B in Table 5 show that a higher share of Black (respectively, Hispanic) admissions for a nursing home predicts higher Black (Hispanic) admissions in the following year as well. The inclusion nursing home fixed effects and the relatively short panel (with only 10 years of data) may raise concerns about the Nickell bias (1981), so in columns 2 and 3 I estimate specifications based on dynamic panel methods from Anderson and Hsiao (1982) and Arellano and Bond (1991; see Appendix section C for more about the dynamic panel specifications). The results for these estimation are similar to the OLS estimates qualitatively, although the exact magnitude differs across specifications. Finally, as a robustness check, Appendix Table A.6 shows that we obtain the same patterns when using numbers of minority residents instead of shares.

As a second test of the equilibrium switch behavior predicted by models of in-group preferences, I estimate event studies of the share of minority residents admitted in a given year, using large positive

[15] Specifically, the model predicts an unstable equilibrium for the minority share in a neighborhood, and a perturbation in the minority share above (respectively, below) this point may lead to a stable equilibrium with the neighborhood being filled mainly with minority (majority) individuals.

shocks to the share of minority residents as the event. Specifically, I estimate event studies of the form:

$$s_{jt}^r = \alpha^{r,shock} + \sum_{k \in \{-F,L\} \backslash \{-1\}} \beta_k^{r,shock} \mathbb{I}[t - E_j^r = k] + \delta_j^{r,shock} + \gamma_t^{r,shock} + \epsilon_{jt}^{r,shock},$$

where E_j^r is the year in which nursing home j receives a much higher than usual number of minority residents of race r, i.e., $s_{j,E_j^r}^r - s_{j,E_j^r-1}^r \geq C$ for some threshold C. Mechanically, $\beta_0^{r,shock}$ will be large given the way the event is defined. However, the real test of in-group preferences is whether this shock leads to persistently higher shares of minority admission in future years, i.e., whether $\beta_k^{r,shock} > 0$ for $k > 0$.[16]

Figures 7a and 7b show results from these event studies using methods from Borusyak, Jaravel, and Spiess (2021), where I consider a year-to-year increase of at least 25pp. in the Black and Hispanic shares of admissions respectively. Consistent with in-group preferences, we observe that a positive shock to the Black (Hispanic) share of admissions of at least 25pp. leads to a more than 10pp. higher share of Black (Hispanic) share of admissions in each of the following 5 years. We observe that there is little evidence of pretrends for Black admissions other than the dip at $t = -1$ (which is mechanical due to how the event is defined), and while some of the pretrend coefficients are statistically significant for Hispanic admissions, the magnitude of these coefficients are far smaller than the effect size. In Appendix Figure A.11, I conduct the same exercise but define the event as either a 10, 15, or 20pp. increase in the Black or Hispanic share of admissions as the "event". We observe qualitatively similar results, with a shock to minority share of admissions having persistent effects, and relatively little pretrends.

Finally, an issue with these tests of whether a shock to the minority share leads to an "equilibrium switch" is that the same pattern can also be explained by an unobserved and persistent change on the supply side. For example, if a new nursing home administrator takes over and decides that the nursing home should admit more (or fewer) minority residents moving forward, the minority share of admissions will also be persistently higher (lower) in the future. Hence, I also conduct a test of in-group preferences, leveraging only within-month variation in occupancy of different races at a nursing home.

Specifically, I run regressions at the nursing home-day level, regressing a measure of new admissions of a particular race on a day on a measure of the racial composition of the nursing home at the end of

[16] Note that I use share of minority admissions in the previous year rather than share of minority residents in the nursing home as the outcome, so that any persistent increase after the shock is not mechanical. In particular, if I instead used share of minority residents in the nursing home as the outcome, then part of the estimated effect in subsequent years will be mechanical if some residents admitted in previous period(s) remain the nursing home for a long duration of time.

the previous day, controlling for nursing home-month fixed effects:

$$admissions^{r}_{jd} = \alpha^{shortrun}_{r} + \sum_{r'} \beta^{shortrun}_{r,r'} occupancy^{r'}_{jd} + \delta^{shortrun}_{r,jm} + \epsilon^{shortrun}_{rjd}.$$

A model of in-group preferences predicts that the coefficient on occupancy of the same race should be positive ($\beta^{shortrun}_{r,r} > 0$), whereas coefficients on occupancy of other races should either be zero or negative if residents dislike other races ($\beta^{shortrun}_{r,r'} \leq 0$ if $r \neq r'$). In the example with the new nursing home administrator, we will have $\beta^{shortrun}_{r,r} = 0$ since the effect will be absorbed by the nursing home-month fixed effects.

The results in Table 6 show that on days when occupancy of a particular race at a nursing home is temporarily elevated, a new resident of the same race is more likely than usual to be admitted. By contrast, a new resident of a different race is either as likely as usual or less likely than usual to be admitted, consistent with in-group preferences. Appendix Tables A.7, A.8, and A.9 show that these results are robust to different definitions of new admissions and occupancies.

3.4 Heterogeneity in Sensitivity to Quality May Also Explain Disparities

While in-group preferences may explain how racial disparities can persist or even worsen over time, this explanation does not shed light on how these racial disparities arose in the first place. One possibility is that minorities may be less sensitive to nursing home quality than Whites. So, to test whether this is plausible, I check an increase (or decrease) in nursing home quality predicts the race of residents admitted in the future. Specifically, I estimate the regression equation:

$$race_i = \alpha^{fe}_{0,r} + \beta^{fe}_{1,r} q_{j(i)} + \beta^{fe}_{2,r} s^{r}_{j(i),t(i)-1} + \delta^{fe}_{j(i),r} + \gamma^{fe}_{c(i),t(i),r} + zip_{i,r} + \epsilon^{fe}_{i,r}, \tag{3}$$

separately for Whites, Blacks and Hispanics, where $race_i$ is a dummy for whether the admitted resident is of a particular race, $q_{j(i)}$ is the quality of the nursing home that she goes to. In the richest specification, I include $\beta^{fe}_{2,r} s^{r}_{j(i),t(i)-1}$ to control for in-group preferences, fixed effects for nursing homes $\delta^{fe}_{j(i),r}$ to control for time-invariant unobserved race-specific preferences for different nursing homes, county-year fixed effects $\gamma^{fe}_{c(i),t(i),r}$ to control for demographic trends, and zip code fixed effects $zip_{i,r}$ to control for residential segregation.

Table 7 shows estimates of equation (3), where I use the quality index as the measure of quality. Comparing the results for different races in different panels, we see that future admissions are more likely to be White and less likely to be Black following an increase in nursing home quality, evidence consistent with White residents being more sensitive to nursing home quality than minority residents. Appendix Tables A.10, A.11, A.12, and A.13 show estimates based on individual quality measures, and the results are qualitatively similar, although we see that these patterns of heterogeneity in sensitivity to quality seem more pronounced for deficiency measures, and less for staffing levels.

A fundamental difficulty with the reduced form evidence on potential causes of racial segregation and disparities presented in this section is that it is difficult to disentangle demand and supply side explanations. For example, following an increase in quality at a nursing home, the increase in probability of new admissions being White is not necessarily due to heterogeneity in sensitivity to quality if the nursing home practices selective admissions. In particular, even if demand for the nursing home may increase proportionally for different races, this will lead to capacity strain, resulting in the nursing home rejecting minority residents at higher rates, giving rise to the same pattern. Hence, in the next section, I introduce a structural model to more cleanly disentangle the different mechanisms.

4 Structural Estimation

4.1 Overview of Empirical Matching Model

In order to disentangle residents' preferences from nursing homes' admission decisions, I estimate an empirical matching model similar to Agarwal and Somaini (2022) and Cheng (2023). I assume that resident i's indirect decision utility for each nursing home $j \in \mathcal{J}_i \equiv \{j | dist_{ij} \leq 15 \; miles\}$ is given by:

$$v_{ij} = w_j' \kappa^1 + w_j' \kappa^2 x_i + dist_{ij}' \kappa^{dist} + \epsilon_{ij}, \tag{4}$$

where x_i and w_j are resident and nursing home characteristics respectively, $dist_{ij}$ is a measure of distance between resident i and nursing home j, and ϵ_{ij} is an idiosyncratic utility shock. For the location normalization, I omit the constant term, and to set the scale normalization, I assume that $\epsilon_{ij} \sim N(0, 1)$.

Nursing homes' admissions policies are given by:

$$\pi_{ij} = x_i' \psi^1 + w_j' \psi^2 x_i + occ_{d(i)j}' \psi^{occ} + \omega_{ij}, \tag{5}$$

where $occ_{d(i)j}$ is a measure of nursing home j's occupancy in the period leading up to i's admission date $d(i)$, and ω_{ij} is an idiosyncratic shock. I assume that nursing home j is willing to admit resident i if and only if $\pi_{ij} \geq \underline{\pi}$, so resident i's constrained choice set is $\{j \in \mathcal{J}_i | \pi_{ij} \geq \underline{\pi}\}$. Note that these constraints are not observed in the data, an issue that the empirical matching model addresses. The location normalization is set by including a constant term in equation (5) and setting $\underline{\pi} = 0$, and the scale normalization is achieved by assuming $\omega_{ij} \sim N(0,1)$.

To elaborate on how this model incorporates elements such as in-group preferences, racial heterogeneity in sensitivity to quality, and discrimination, equation (4) typically takes the form:

$$\begin{aligned} v_{ij} =& \kappa_0^{black} s_{d(i)j}^{black} + \kappa_0^{hisp} s_{d(i)j}^{hisp} + \kappa_1^{black} s_{d(i)j}^{black} black_i + \kappa_1^{hisp} s_{d(i)j}^{hisp} hispanic_i \\ &+ q_j' \kappa_0^q + black_i q_j' \kappa_{black}^q + hispanic_i q_j' \kappa_{hisp}^q + dist_{ij}' \kappa^{dist} + \epsilon_{ij}, \end{aligned}$$

where $s_{d(i)j}^r$ is the share of nursing home j's admissions that are of race r in the 365 days leading up to i's admission date $d(i)$. We can think of κ_0^r and $\kappa_0^r + \kappa_1^r$ respectively as measuring preferences among those of race $r' \neq r$ and race r for a higher share of recently admitted residents being of race r. In the absence of racial preferences among residents, we will expect $\kappa_0^r = \kappa_1^r = 0$. Similarly, κ_0^q captures White residents' demand for quality, whereas $\kappa_0^q + \kappa_r^q$ captures demand for quality among residents of race r (for non-White residents). If there was no racial heterogeneity in sensitivity to quality, then we should have $\kappa_r^q = 0$. In the simplest specification for the supply side, I estimate:

$$\pi_{ij} = \psi_0 + \psi^{black} black_i + \psi^{hisp} hispanic_i + \tilde{x}_i' \psi^{\tilde{x}} + q_j \psi^q + occ_{d(i)j}' \psi^{occ} + \omega_{ij},$$

where $\tilde{x}_i$ are some none-race characteristics of resident i. In the absence of discriminatory admissions practices, we would expect $\psi^r = 0$ for both minority races.

Agarwal and Somaini (2022) provide a sharp set of identification conditions for such a model, and the key substantive requirement is that we need both demand and supply side instruments. Hence, I use distance as the demand side instrument and temporary fluctuations in nursing homes' occupancy (specifically log of lagged seven-day occupancy residualized of nursing home-month fixed effects) as the supply instrument. The relevance condition for both instruments are likely to be satisfied: residents have a strong preference for nursing homes close to where they used to live, and nursing homes are less likely to admit new residents when capacity is strained, as seen in Table 4a. The exclusion restriction for the demand instrument is also likely satisfied, since nursing homes are unlikely to be care about where their residents originally lived.

The exclusion restriction for the supply instrument deserves closer scrutiny, and for better intuition we start by considering why using occupancy (instead of temporary fluctuations in occupancy) is likely to violate the exclusion restriction, and its implications for structural model's estimates. All else equal, residents may prefer less crowded nursing homes (thus violating the exclusion restriction), and if higher-quality nursing homes are in greater demand, this will lead us to underestimate residents' demand for quality. Moreover, if preferences for "crowdedness" also varies by race, then our estimates of heterogeneity in sensitivity to quality will also be biased.

The use of *temporary fluctuations* in nursing home occupancy (specifically, within nursing home-month fluctuations) addresses this issues in two ways. First, short-term fluctuations in occupancy are less likely to matter for residents. Second, by residualizing the occupancy measure of nursing home-month fixed effects, we eliminate potential correlations between occupancy and nursing home characteristics such as quality, as well as the share minority in the past year. Figure 7 illustrates this by showing that the distribution of temporary occupancy fluctuations of nursing homes within 15 miles of each resident is essentially identical across above-median and below-median-quality nursing homes close to White, Black, and Hispanic residents. Similarly, Figure 8 shows that the distribution of the supply instrument is identical when we split by above- and below-median share of Black and Hispanic admissions in the past year. Finally, Appendix Figures A.12, A.13, A.14, A.15, and A.16 show that the same pattern holds if we consider individual quality measures, and Appendix Figure A.17 shows that the same results hold when we interact the share of admissions Black or Hispanic with the race of the resident.

Finally, estimation of this model has to deal with the curse of dimensionality. In particular, there are $2^{|\mathcal{J}_i|} - 1$ possible constrained choice sets for resident i. The average number of nursing homes $|\mathcal{J}_i|$ within 15 miles of a resident in California is 50 for example, and can be as large 200 for some residents. Hence, methods such as maximum likelihood that require us to sum over each distinct possibility are computationally infeasible. Therefore, I use Gibbs sampling with data augmentation (for v_{ij} and π_{ij}) for my estimation, since this obviates the need to individually compute the probability of each potential choice set, without needing to make additional substantive assumptions.

At a high level, in each iteration of the Gibbs sampler, I draw utility and profit shocks ϵ_{ij} and ω_{ij} in such a way that the resulting latent variables respect the matching outcomes. I then update the posterior distribution of the parameters before moving onto the next iteration. For example, if resident i is admitted to a nursing home $\mu(i)$, when drawing the utility shock $\epsilon_{i,\mu(i)}$ we must make sure that the resulting indirect utility $v_{i,\mu(i)}$ must be no smaller than v_{ij} for any $j \in \mathcal{J}_i$ where $\pi_{ij} \geq 0$, so we draw $\epsilon_{i,\mu(i)}$ from a distribution that is truncated from below (for details on the algorithm for the Gibbs sampler, see Appendix section D). Under regularity conditions, the draws of the parameters will eventually converge to their stationary distribution. We can then form Bayesian confidence sets based on the distribution of these draws, which are also endowed with a frequentist interpretation as a consequence of the Bernstein von-Mises theorem. Due to computational constraints, I limit the sample to residents and nursing homes in four different states between 2008–2010.

4.2 Structural Estimation Results

Table 8 shows results from my estimation of the empirical matching model, separately for the four largest states in my sample that had a decent share of Black and Hispanic residents. We observe that the coefficient estimates for the demand and supply instruments (distance and occupancy respectively) have the expected sign and are highly statistically significant, which is reassuring for the identification of the model.

Starting with the demand estimates, we observe that the coefficient estimates on the interactions between the resident's own race and the share of nursing home admissions being of the same race in the past year are positive and highly statistically significant, consistent with in-group preferences. In addition, the estimated coefficients for the interaction terms between quality and the resident being Black or Hispanic are negative, and are mostly statistically significant, suggesting that minority residents are less sensitive to nursing home quality than White residents. Turning to the supply side, we see that the coefficient estimates on Black and Hispanic are negative and highly statistically significant, consistent with discrimination against minorities in nursing homes' admissions policies.

To account for time-invariant preferences for nursing homes that are not captured by the variables included in the utility equation, in Appendix Table A.14 I include nursing home fixed effects in the utility equation as a robustness check. We see that the estimates associated with in-group preferences, lower sensitivity to quality among minorities, and discrimination against minorities remain statistically significant with the inclusion of these fixed effects in utility. In addition, Appendix Table A.15 shows that these results are robust to the use of multiple quality measures (namely RN staffing, LPN staffing, standard deficiencies, and complaint deficiencies) instead of the quality index, with minority residents generally being less sensitive to the individual quality measures than White residents as well.

While in-group preferences, heterogeneity in sensitivity to quality, and selective admissions practices along racial dimensions seem to be present in all four states, it is difficult to compare the importance of each of these explanations (and residential segregation) in explaining segregation and disparities based on the coefficients alone. Hence, I address this in the next section using counterfactual simulations.

5 Counterfactuals

In this section, I conduct counterfactual simulations to assess how important different factors are for explaining segregation and choice disparities. Specifically, I use the structural estimates from the previous section to simulate the dynamic evolution of segregation and disparities without change. by modifying different parameters. In these simulations, I abstract away from endogenous quality adjustments by nursing homes, setting quality for each nursing home to its average over the time period of the structural estimation (2008–2010).

To simulate a successful ban on discriminatory admissions practices nursing homes, I set the race-specific parameters ψ^r in the admissions equation to zero. To mimic the elimination of in-group preferences, I set the parameters for race $(\kappa_0^{r\prime}, \kappa_1^{r\prime})'$ in residents' utility equation to zero. Next, to simulate the homogeneous sensitivity to quality, I set the interactions between nursing home quality and minority race dummies $\kappa_r^q = 0$. Finally, to mimic the elimination of residential segregation, I randomize the zip code of prior residence for each resident, effectively randomizing them to different counterfactual choice sets. For more details on the simulations, see Appendix section E.

To measure the effect of each explanation on segregation, I compute the average statewide dissimilarity index over time under the relevant counterfactual. I then compare it to the average statewide similarity index in the "status-quo" simulations which use the estimated structural parameters without modification, and plot the average percent reduction in the dissimilarity index over time. I use 5000 days for the simulations and 100 replications of each simulation. Given that the initial distribution of residents across nursing homes in these simulations are somewhat arbitrary, I plot the dissimilarity index starting from day 2500 to allow the estimates to stabilize.

Figure 9 shows the effect of various counterfactual policies on the dissimilarity index for Black residents in the four states. We observe that residential segregation is by far the most important explanation for nursing home segregation in all four states. Appendix Figure A.18 plots the dissimilarity index for the same counterfactuals for Hispanic residents and we see a similar pattern with residential segregation being the largest contributor to nursing home segregation.

I measure the effect of different explanations on racial disparities in a similar way: in the simulations corresponding to each counterfactual (as well as the status-quo simulations), for every 100 days, I estimate the racial gaps based on equation (1) using simulated data from the past 100 days. I then plot the changes in the Black-White gap and Hispanic-White gaps in the counterfactual simulations relative to the status-quo simulations in Figures 10 and 11 respectively.

Four striking facts stand out from these figures. First, in contrast to the results for nursing home segregation, residential segregation is typically not the most important explanation for racial disparities. It is the most important explanation for the Black-White gap only in Texas and New York (where it is tied with discrimination).

Second, eliminating residential segregation can potentially widen disparities for some minorities, as in the case for Hispanics in Florida and Texas. This surprising finding can be explained by the fact that the average quality index of nursing homes close to Hispanic residents is higher than for Whites in these two states, as shown in Appendix Table A.3.

Third, the most important explanations for disparities between minority groups vis-à-vis White residents may be different for different minority groups, even within the same state. In fact, the effect of an explanation may even have opposite signs for different minority groups: we observe that in both Florida and Texas eliminating residential segregation worsens Hispanic-White disparities, but reduces Black-White disparities.

Fourth, the importance of different explanations for reducing disparities differs across states. Each of the four explanations is shown to be most important for explaining Black-White disparities in at least one state. On the other hand, heterogeneity in sensitivity to quality explains most of the Hispanic-White gap in three of the four states (California, Texas, and Florida, but not New York).

In the final exercise for the simulations, we study how the different factors potentially interact with each other. Specifically, we progressively add more explanations to the counterfactual, starting with only the elimination of residential segregation, and then adding the elimination of in-group preferences, selective admissions, and finally heterogeneity in sensitivity to quality.

Figure 12 shows the effect on residential segregation when we eliminate one or more factors simultaneously. While we previously saw that residential segregation was by far the most important explanation for nursing home segregation when the factors when considered individually, some of the other explanations now play a much larger role when considered in addition to residential segregation. In particular, we see that eliminating in-group preferences and selective admissions further reduce segregation, while eliminating heterogeneity in sensitivity to quality does not seem to affect segregation.

Moreover, there seems to be synergism between these explanations. For example, eliminating selective admissions by itself has mixed effects, reducing segregation by less than 20 percent in the best case and worsening segregation in others, whereas it now reduces segregation consistently by 20–30 percent when residential segregation and in-group preferences have already been eliminated.

Finally, Figures 13 and 14 show the effect of eliminating multiple factors on disparities. We see that in California, Florida, and New York, we see that disparities for both Blacks and Hispanics (relative to Whites) fall progressively for the most part as we eliminate more factors, as we would expect. Texas is the only outlier, where eliminating selective admissions and heterogeneity in sensitivity to quality (on top of residential segregation and in-group preferences) seems to worsen disparities slightly.

6 Conclusion

In this paper, I studied the question of why racial minorities are disproportionately concentrated in low-quality nursing homes. I find evidence that residential segregation, in-group preferences, discrimination, and heterogeneity in sensitivity to quality may all play a role. In terms of magnitude, residential segregation seems to be the most important contribution to nursing home segregation, but this is not so for racial disparities. Instead, the relative importance of different explanations for racial disparities vary widely across different states. Moreover, eliminating residential segregation may be detrimental for the Hispanic-White gap in states such as Florida, although it also reduces the Black-White gap in this case.

These findings are relevant to policymaking in several ways. First, policymakers need to think carefully about whether their goal is to reduce segregation or to reduce disparities, since the most effective solutions for the two goals may be different. Second, the most effective policies for reducing disparities may depend greatly on local conditions. Third, there is a possibility that policies that reduce disparities for one minority group may not work for other minority groups (or may even have the opposite effect).

References

[1] Aaronson, Daniel, Daniel Hartley, and Bhashkar Mazumder, 2021. "The effects of the 1930s HOLC "redlining" maps." *American Economic Journal: Economic Policy*, 13(4): 355–392.

[2] Agarwal and Somaini, 2022. "Demand Analysis under Latent Choice Constraints." Working paper.

[3] Anderson, Theodore Wilbur, and Cheng Hsiao, 1982. "Formulation and estimation of dynamic models using panel data." *Journal of Econometrics*, 18(2): 47–82.

[4] Arellano, Manuel, and Stephen Bond, 1991. "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations." *The Review of Economic Studies*, 58(2): 277–297.

[5] Arrow, Kenneth J. 1972a. "Models of Job Discrimination." In *Racial Discrimination in Economic Life*, edited by Anthony H. Pascal, 83–102. Lexington, MA: D.C. Heath.

[6] Arrow, Kenneth J. 1972b. "Some Mathematical Models of Race Discrimination in the Labor Market." In *Racial Discrimination in Economic Life*, edited by Anthony H. Pascal, 187–204. Lexington, MA: D.C. Heath.

[7] Baicker, Katherine, and Amitabh Chandra, 2004. "Medicare Spending, The Physician Workforce, And Beneficiaries' Quality Of Care: Areas with a high concentration of specialists also show higher spending and less use of high-quality, effective care." *Health Affairs*, 23(Suppl1): W4–184.

[8] Becker, Gary S., (1957) 1971. *The Economics of Discrimination.* 2nd ed. Chicago: Chicago University Press.

[9] Billings, Stephen B., David J. Deming, and Jonah Rockoff, 2014. "School segregation, educational attainment, and crime: Evidence from the end of busing in Charlotte-Mecklenburg." *The Quarterly Journal of Economics*, 129(1): 435–476.

[10] Black, Sandra E, 1999. "Do better schools matter? Parental valuation of elementary education." *The Quarterly Journal of Economics*, 114(2): 577–599.

[11] Borusyak, Kirill, Xavier Jaravel, and Jann Spiess, 2022. "Revisiting event study designs: Robust and efficient estimation." Working paper.

[12] Card, David, Alexandre Mas, and Jesse Rothstein, 2008. “Tipping and the Dynamics of Segregation.” *The Quarterly Journal of Economics*, 123(1): 177–218.

[13] Cheng, Alden, 2023. “Demand for Quality in the Presence of Information Frictions: Evidence from the Nursing Home Market.” Working paper.

[14] Christensen, Peter, and Christopher Timmons, 2023. “The Damages and Distortions from Discrimination in the Rental Housing Market.” *The Quarterly Journal of Economics*, 138(4): 2505–2557.

[15] Derenoncourt, Ellora, 2022. “Can you move to opportunity? Evidence from the Great Migration.” *American Economic Review*, 112(2): 369–408.

[16] Gandhi, Ashvin, 2020. “Picking your patients: Selective admissions in the nursing home industry.” Working paper.

[17] Hackmann, Martin B., 2019. “Incentivizing better quality of care: The role of Medicaid and competition in the nursing home industry.” *American Economic Review*, 109(5): 1684–1716.

[18] Hailey, Chantal A., 2022. “Racial preferences for schools: Evidence from an experiment with White, Black, Latinx, and Asian parents and students.” *Sociology of Education*, 95(2): 110–132.

[19] Hurd, Michael D., Pierre-Carl Michaud, and Susann Rohwedder, 2017. “Distribution of lifetime nursing home use and of out-of-pocket spending.” *PNAS*, 104(37): 9838–9842.

[20] Jain, Vardhmaan, Mahmoud Al Rifai, Michelle T. Lee, Ankur Kalra, Laura A. Petersen, Elizabeth M. Vaughan, Nathan D. Wong, Christie M. Ballantyne, and Salim S. Virani, 2021. “Racial and geographic disparities in internet use in the US among patients with hypertension or diabetes: implications for telehealth in the era of COVID-19.” *Diabetes Care*, 44(1): e15.

[21] Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz, 2007. “Experimental analysis of neighborhood effects.” *Econometrica*, 75(1): 83–119.

[22] Li, Yue, Charlene Harrington, Helena Temkin-Greener, Kai You, Xueya Cai, Xi Cen, and Dana B. Mukamel, 2015. “Deficiencies in care at nursing homes and racial/ethnic disparities across homes fell, 2006–11.” *Health Affairs*, 34(7): 1139–1146.

[23] Mack, Deborah S., Bill M. Jesdale, Christine M. Ulbricht, Sarah N. Forrester, Pryce S. Michener, and Kate L. Lapane, 2020. “Racial segregation across US nursing homes: A systematic review of measurement and outcomes.” *The Gerontologist*, 60(3): e218–e231.

[24] Margo, Robert A., 1991. "Segregated schools and the mobility hypothesis: A model of local government discrimination." *The Quarterly Journal of Economics*, 106(1): 61-73.

[25] Nickell, Stephen, 1981. "Biases in dynamic models with fixed effects." Econometrica, 49(6): 1417–1426.

[26] Oosterbeek, Hessel, Sándor Sóvágó, and Bas van der Klaauw, 2021. "Preference heterogeneity and school segregation." *Journal of Public Economics*, 197: 104400.

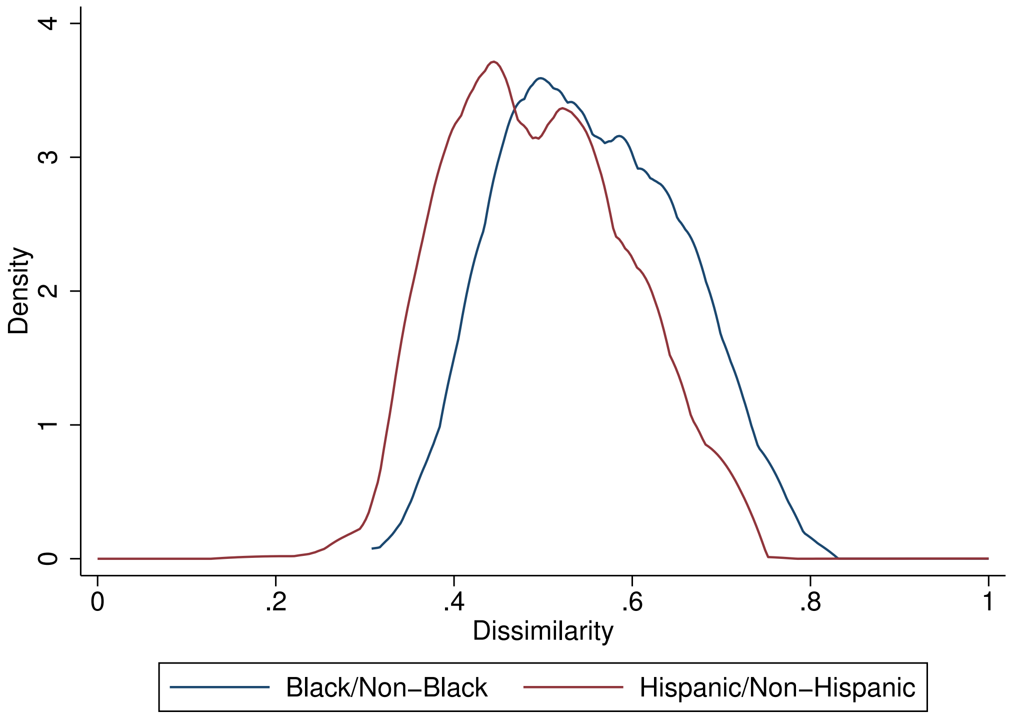
[27] Rahman, Momotazur, and Andrew D. Foster, 2015. "Racial segregation and quality of care disparity in US nursing homes." *Journal of Health Economics*, 39: 1–16.

[28] Schelling, Thomas C., 1971. "Dynamic models of segregation." *Journal of Mathematical Sociology*, 1(2): 143–186.

Appendix

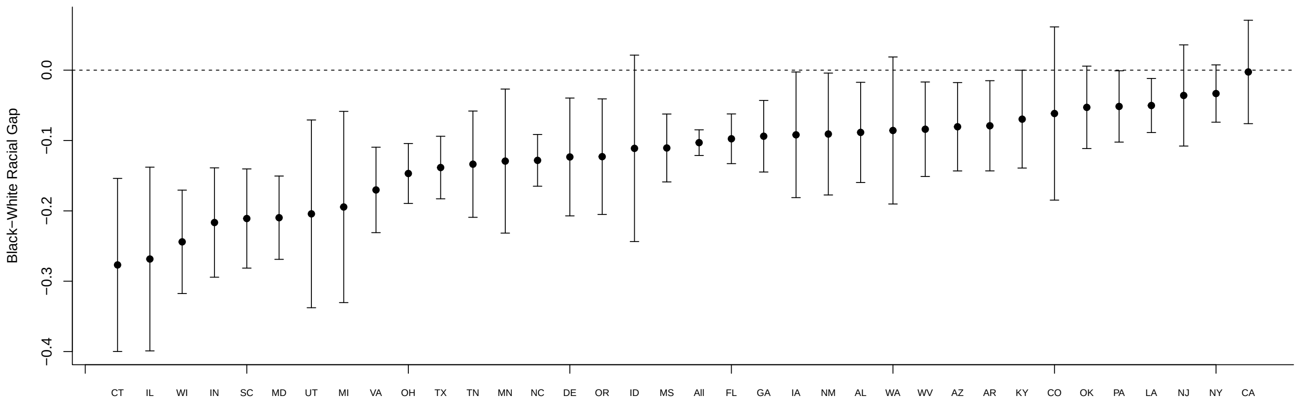
A Tables and Figures

Figure 1: Patterns of Racial Segregation (Statewide Index of Dissimilarity)

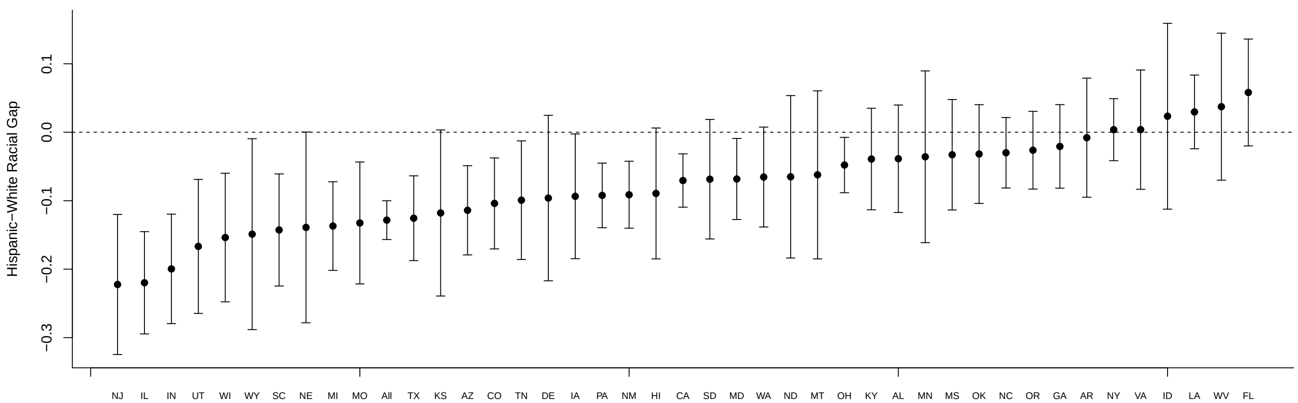


Notes: This figure shows kernel density plots of the dissimilarity index for Black versus non-Black residents and Hispanic versus non-Hispanic residents, measured at the state level.

Figure 2: Racial Gaps in Nursing Home Quality



(a) Black-White Gap

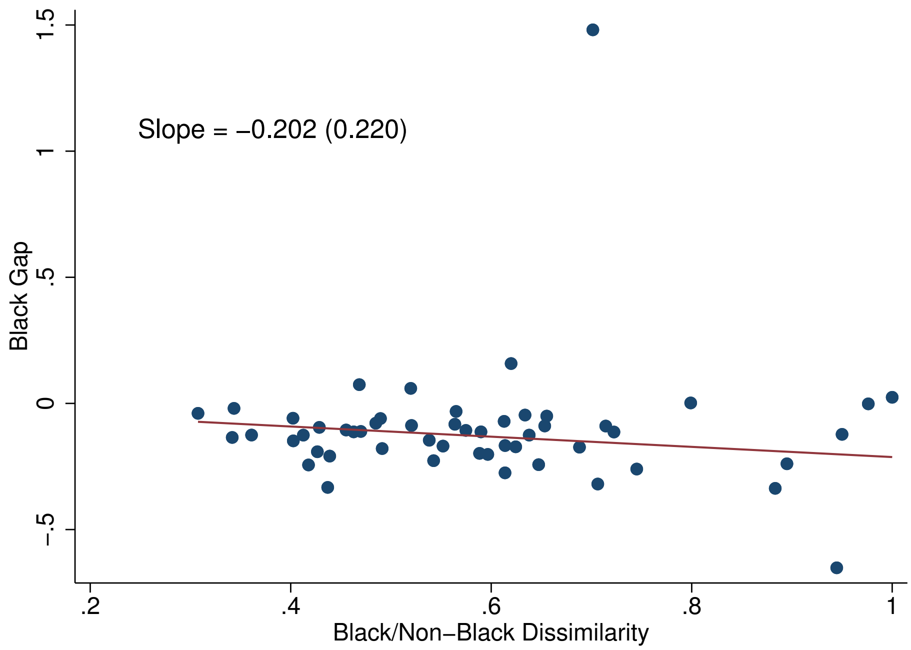


(b) Hispanic-White Gap

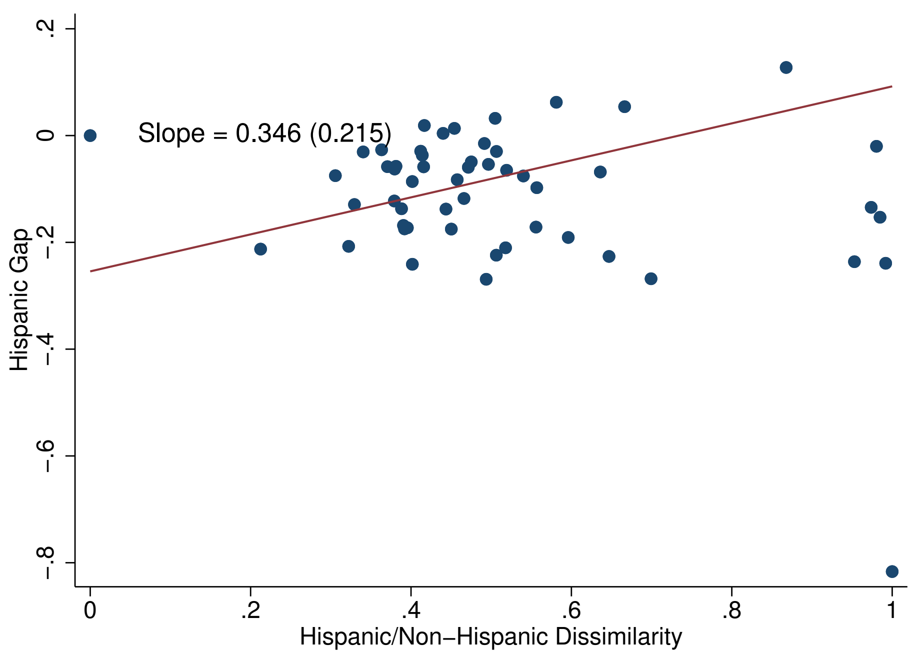
Notes: These figures display the estimated racial gaps in nursing home quality by state. The quality measure is an index constructed by first taking the sum of standardized RN staffing, LPN staffing, and negative standard deficiencies, and then standardizing the sum (weighting by number of residents). Error bars indicate 95 percent confidence intervals for the estimates.

Figure 3: Cross-Sectional Relationship Between State-Level Segregation and Disparities

(a) Black/Non-Black

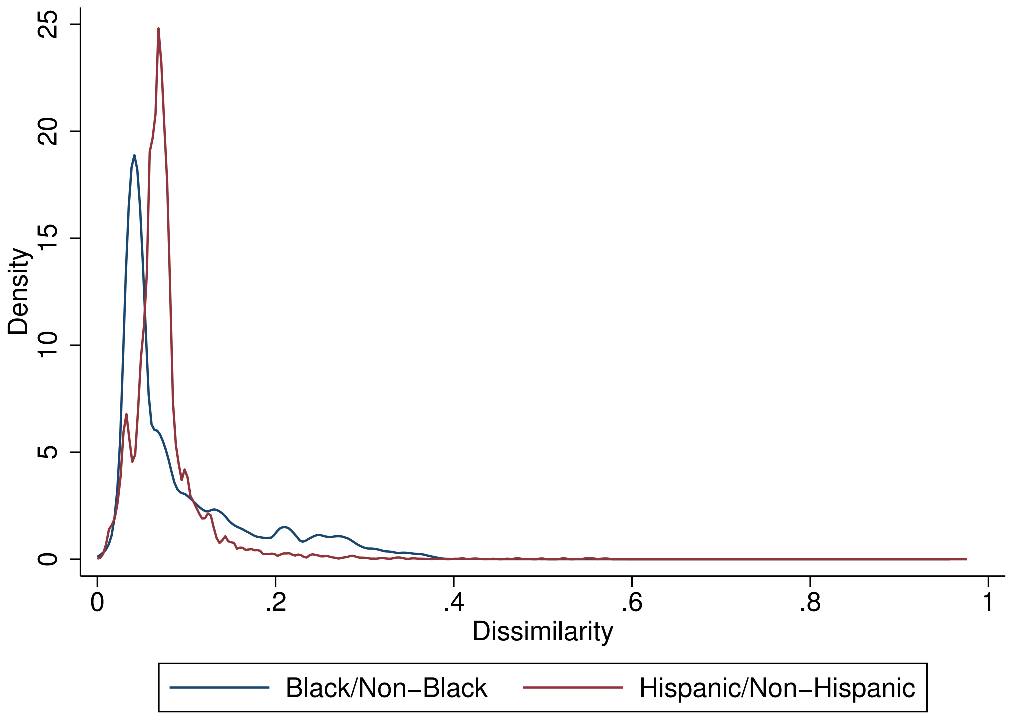


(b) Hispanic/Non-Hispanic



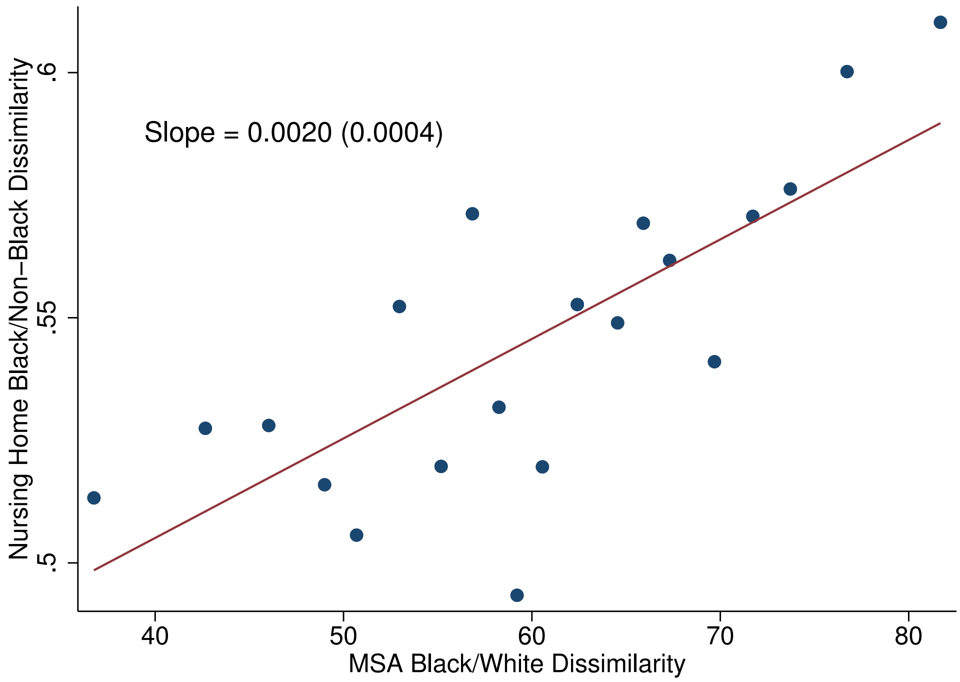
Notes: These figures display scatter plots of the estimated racial gap (based on the quality index) against segregation at the state level. Observations are weighted by the number of residents admitted to the state.

Figure 4: Patterns of Racial Segregation (Index of Dissimilarity Based on 15-Mile Radius)



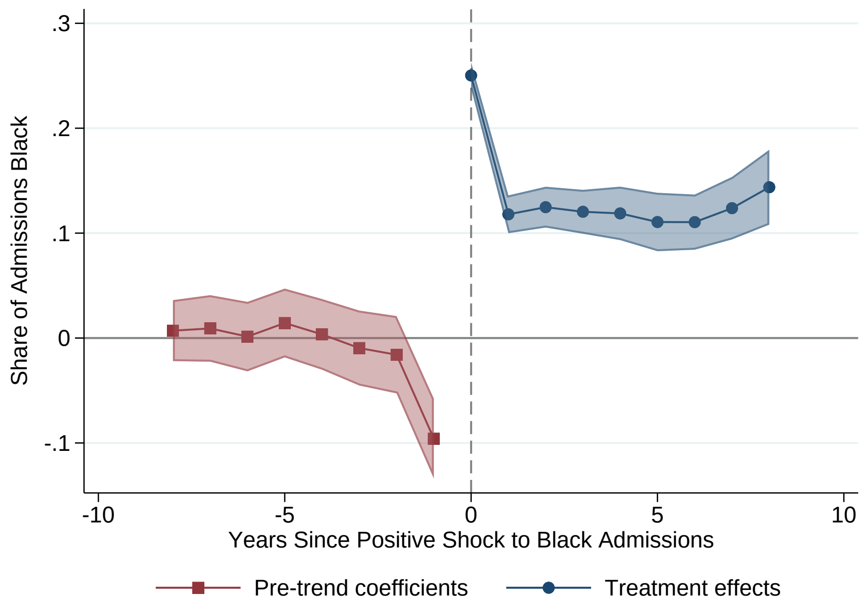
Notes: This figure shows kernel density plots of the dissimilarity index for Black versus non-Black residents and Hispanic versus non-Hispanic residents, measured based on neighborhoods in a 15-mile radius of each zip code and at the state level.

Figure 5: Association Between Residential and Nursing Home Segregation

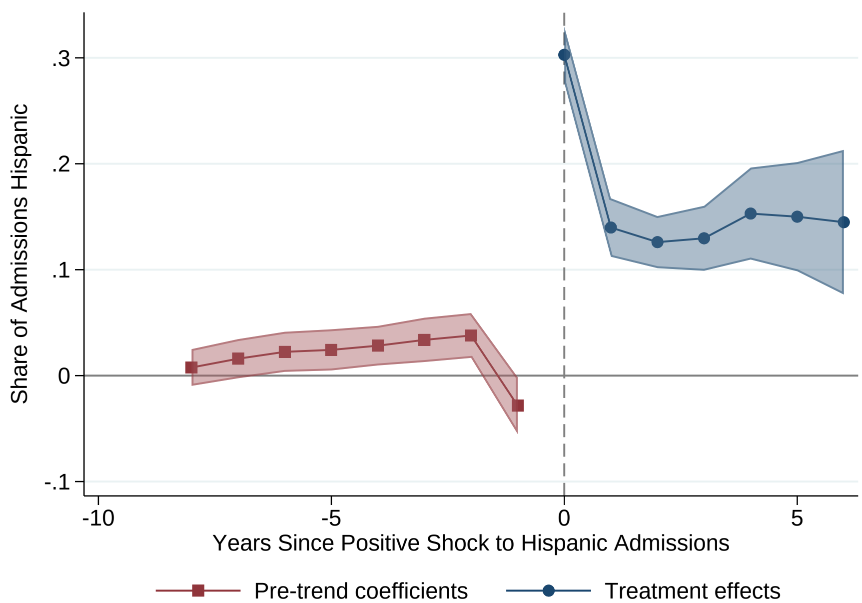


Notes: This figure shows a binscatter of MSA-level Black-White dissimilarity indices for residential segregation against state-level Black/non-Black nursing home dissimilarity indices. For MSAs that straddle multiple states, I create “duplicates” when merging them to state-level nursing home segregation data. Observations are weighted by the number of first stays for nursing home residents in the state during 1999–2010.

Figure 6: Event Study on the Effect of a Positive Shock to the Share of Minority Admissions



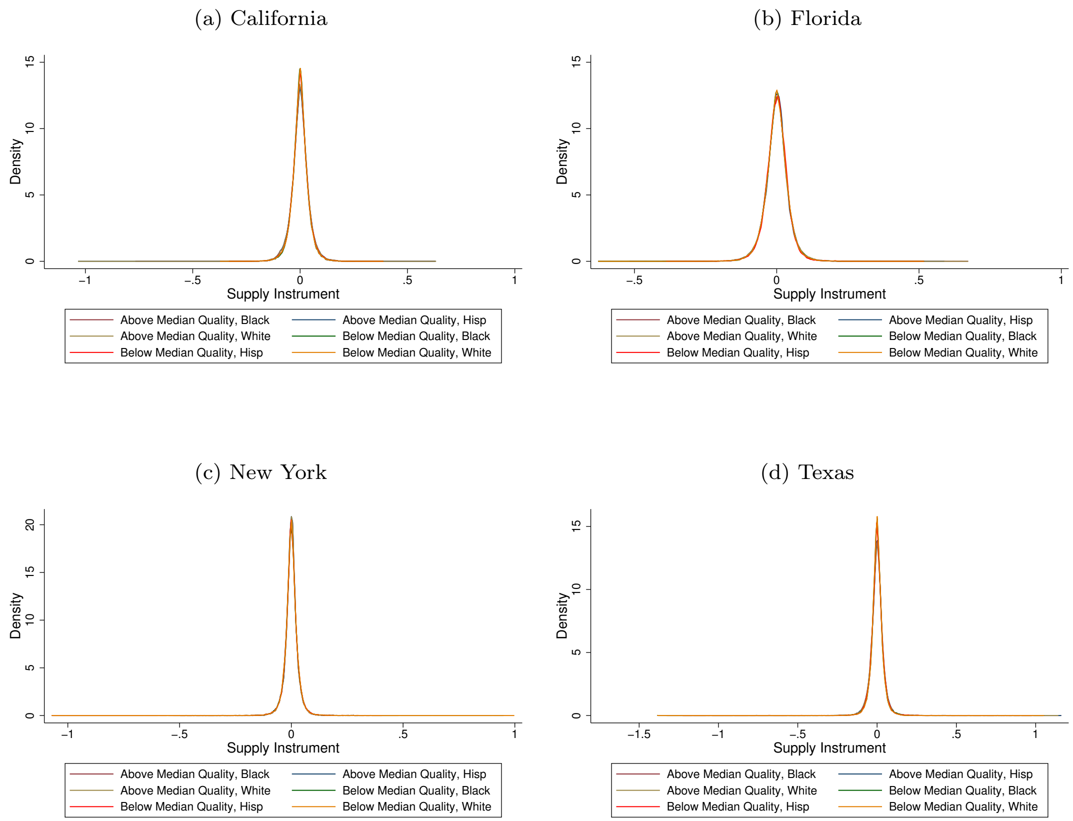
(a) Increase in Share of Black Admissions



(b) Increase in Share of Hispanic Admissions

Notes: These figures show event study estimates of the effect of a large shock to the minority share of admissions (defined as a year-to-year increase in the minority share of admissions of at least 25pp.) on future minority share of admissions, based on event study methods of Borusyak, Jaravel, and Spiess (2021).

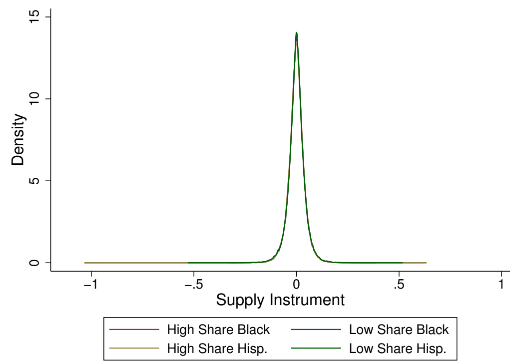
Figure 7: Exclusion Restriction for Temporary Occupancy Fluctuations (Split by Quality Index and Race of Resident)



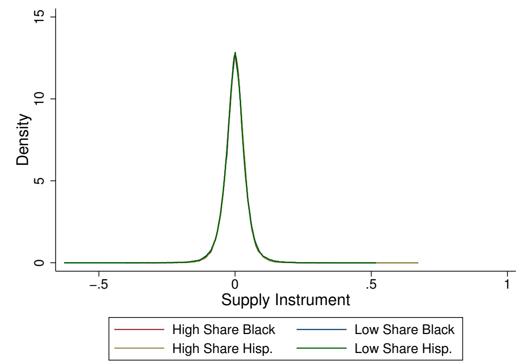
Notes: This figure display kernel density plots of temporary occupancy fluctuations in the four states used for the structural estimation (defined as lagged 7-day log occupancy residualized of nursing home-month fixed effects) of nursing homes within 15 miles of each resident at their time of admission. The unit of observation is a nursing home-resident pair, and the sample is split by whether nursing homes are above-median or below-median in quality (based on the quality index), and the race of the resident.

Figure 8: Exclusion Restriction for Temporary Occupancy Fluctuations (Split by Share of Minority Admissions in the Past Year)

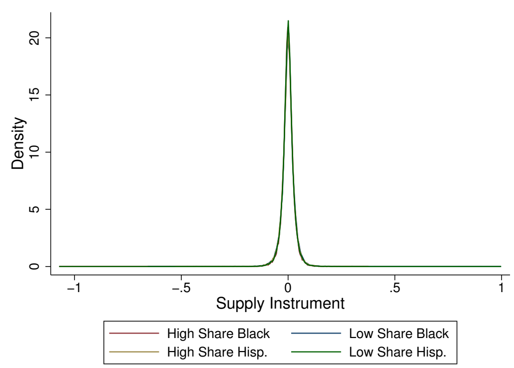
(a) California



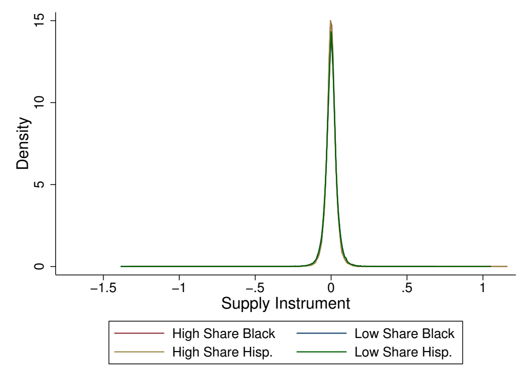
(b) Florida



(c) New York

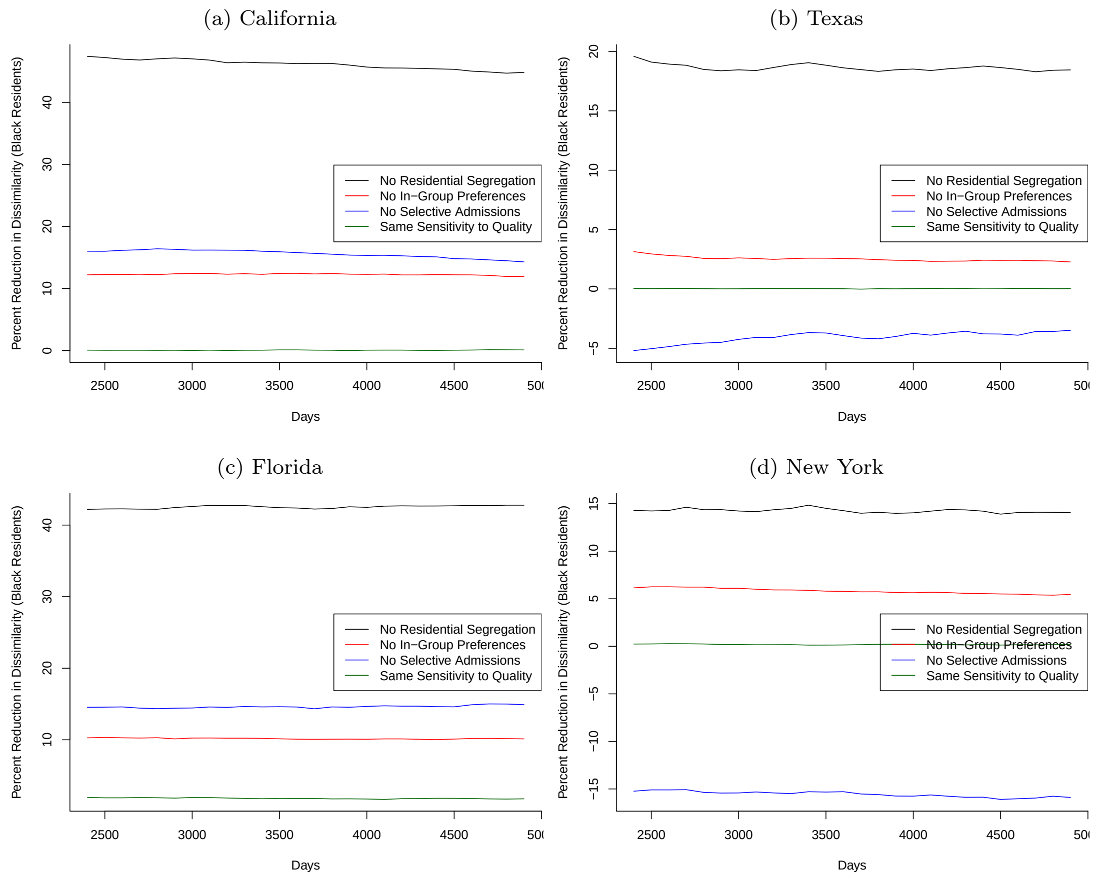


(d) Texas



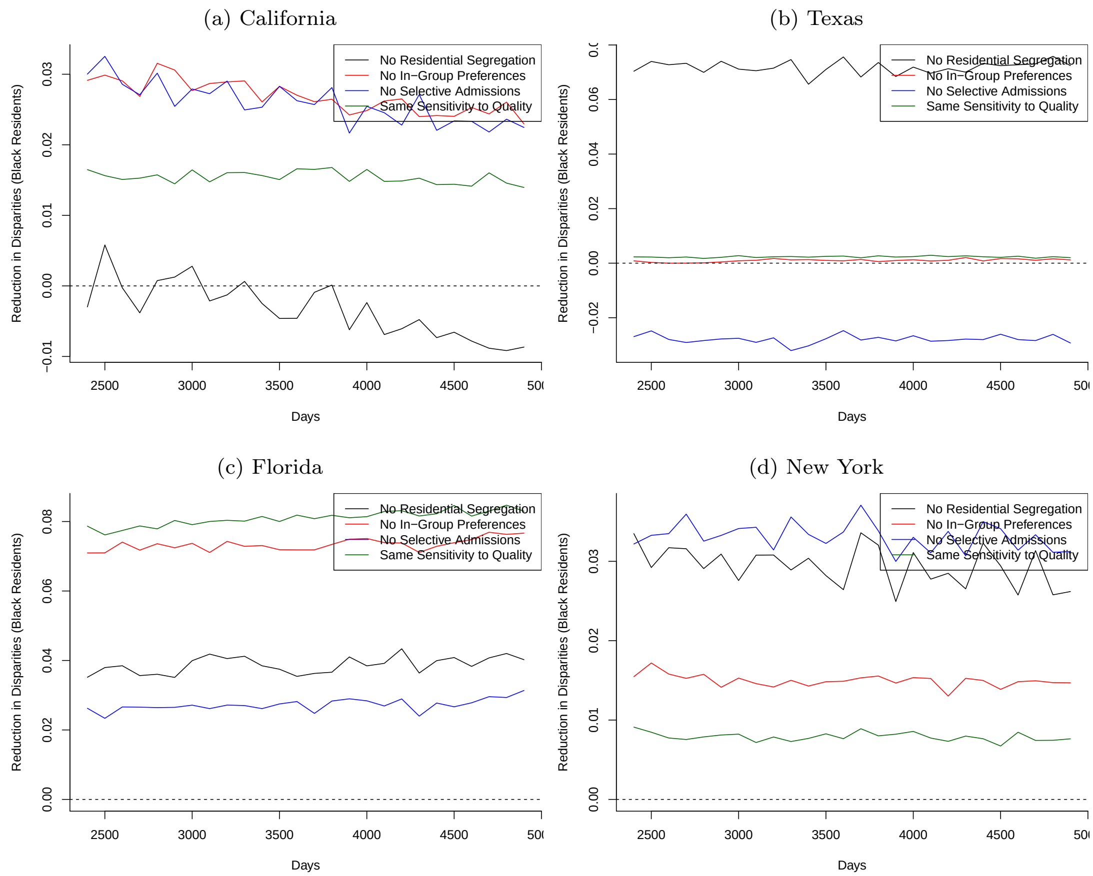
Notes: This figure display kernel density plots of temporary occupancy fluctuations in the four states used for the structural estimation (defined as lagged 7-day log occupancy residualized of nursing home-month fixed effects) of nursing homes within 15 miles of each resident at their time of admission. The unit of observation is a nursing home-resident pair, and the sample is split by whether the share of nursing home admissions Black (or Hispanic) in the last 365 days is above or below the median for this share.

Figure 9: Counterfactual Segregation (Black vs Non-Black Residents)



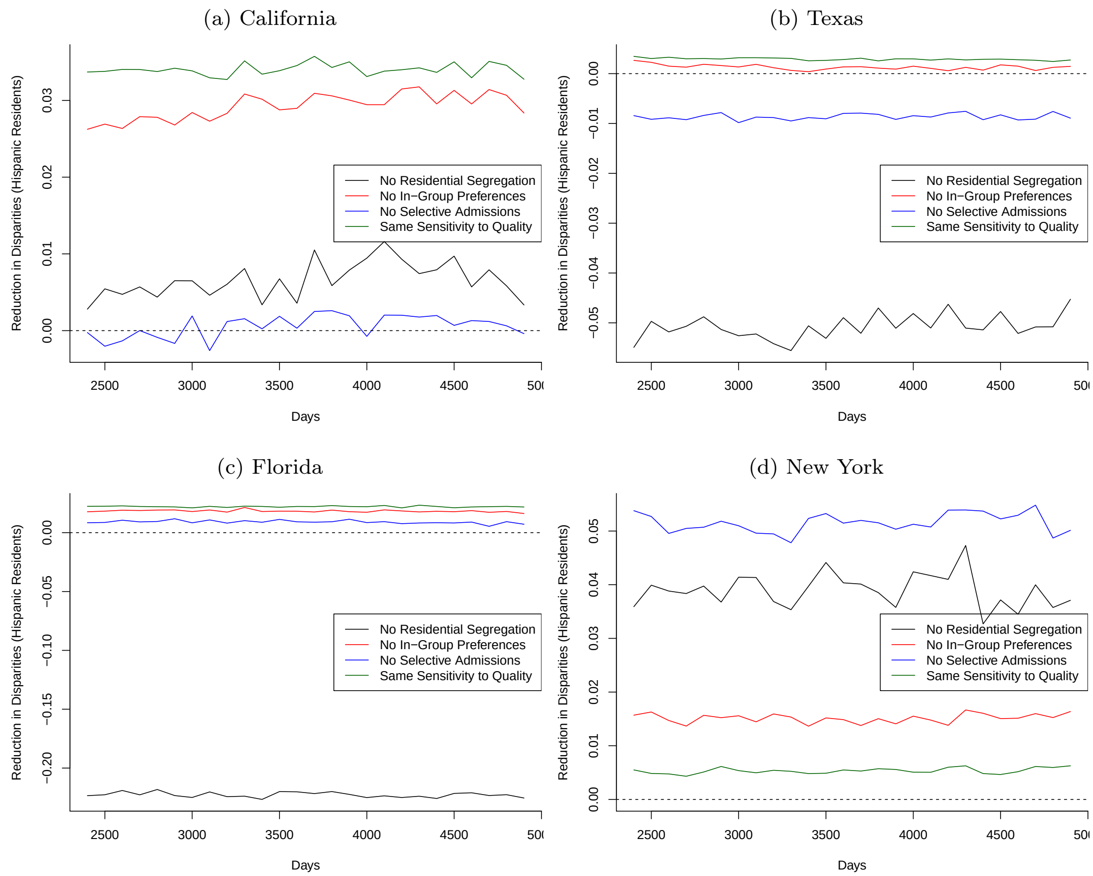
Notes: These figures display the simulated reduction in average Black/non-Black dissimilarity index over time separately for different counterfactuals compared to the status quo simulations, over 100 replications for each counterfactual.

Figure 10: Counterfactual Racial Disparities (Black vs White Residents)



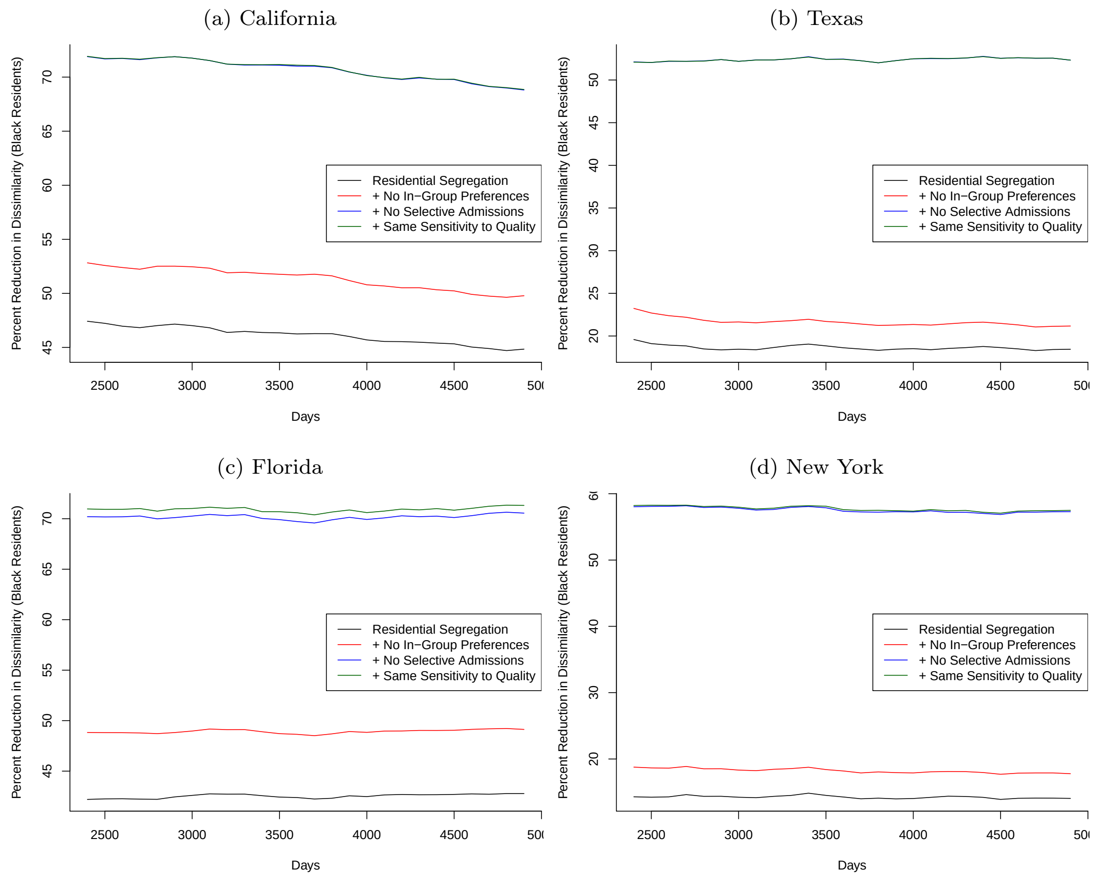
Notes: These figures display the simulated average difference in quality of nursing homes that Blacks versus Whites are admitted to over time separately for different counterfactuals relative to the status quo simulations, over 100 replications for each counterfactual.

Figure 11: Counterfactual Racial Disparities (Hispanic vs White Residents)



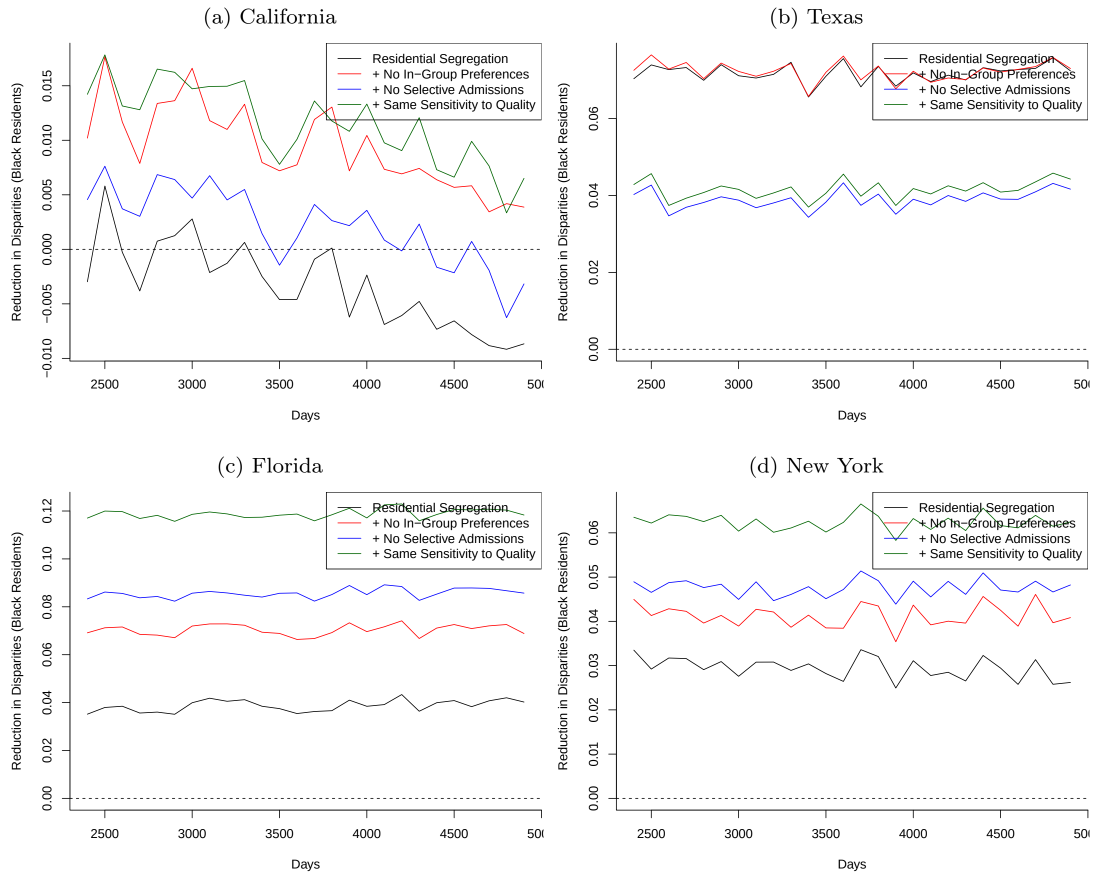
Notes: These figures display the simulated average difference in quality of nursing homes that Hispanics versus Whites are admitted to over time separately for different counterfactuals relative to the status quo simulations, over 100 replications for each counterfactual.

Figure 12: Counterfactual Segregation with Several Explanations (Black vs Non-Black Residents)



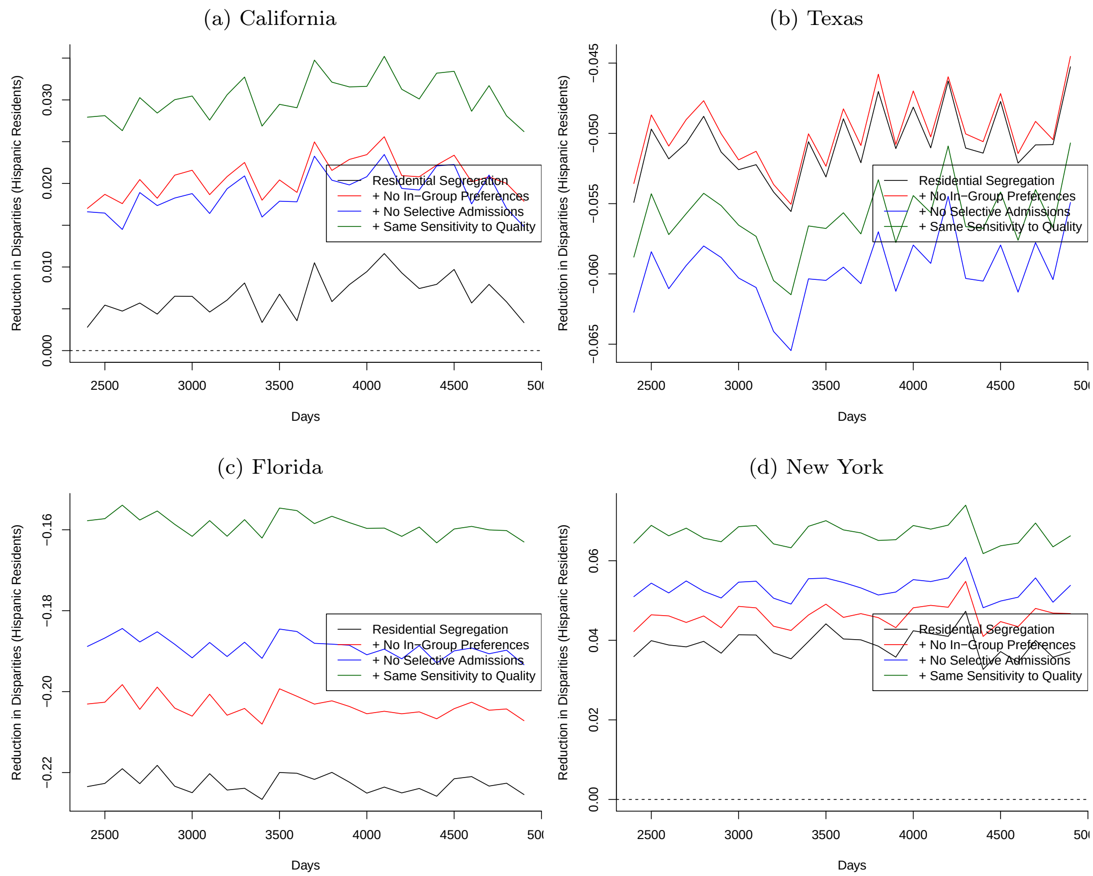
Notes: These figures display the simulated average Black/non-Black dissimilarity index over time separately for different counterfactuals relative to the status quo simulations, over 100 replications for each counterfactual.

Figure 13: Counterfactual Racial Disparities with Several Explanations (Black vs White Residents)



Notes: These figures display the simulated average difference in quality of nursing homes that Blacks versus Whites are admitted to over time separately for different counterfactuals relative to the status quo simulations, over 100 replications for each counterfactual.

Figure 14: Counterfactual Racial Disparities with Several Explanations (Hispanic vs White Residents)



Notes: These figures display the simulated average difference in quality of nursing homes that Hispanics versus Whites are admitted to over time separately for different counterfactuals relative to the status quo simulations, over 100 replications for each counterfactual.

Table 1: Summary Statistics for Residents

	All	White	Black	Hispanic
Age	77.57	78.63	71.15	73.29
	(12.76)	(11.98)	(15.27)	(14.76)
Female	0.623	0.633	0.573	0.555
	(0.485)	(0.482)	(0.495)	(0.497)
Married	0.324	0.335	0.224	0.330
	(0.468)	(0.472)	(0.417)	(0.470)
Less than High School	0.271	0.240	0.409	0.542
	(0.445)	(0.427)	(0.492)	(0.498)
High School/Some College	0.594	0.619	0.499	0.376
	(0.491)	(0.486)	(0.500)	(0.484)
At Least Bachelor's Degree	0.108	0.117	0.0586	0.0416
	(0.310)	(0.321)	(0.235)	(0.200)
Medicare	0.714	0.731	0.644	0.580
	(0.452)	(0.443)	(0.479)	(0.494)
Medicaid	0.0991	0.0768	0.208	0.217
	(0.299)	(0.266)	(0.406)	(0.413)
Admitted from Acute Care Hospital	0.843	0.841	0.850	0.856
	(0.363)	(0.365)	(0.357)	(0.351)
Admitted from Home	0.110	0.113	0.0964	0.0989
	(0.313)	(0.317)	(0.295)	(0.299)
Dementia	0.246	0.248	0.245	0.231
	(0.431)	(0.432)	(0.430)	(0.421)
White	0.833			
	(0.373)			
Black	0.102			
	(0.302)			
Hispanic	0.0397			
	(0.195)			
Number of Residents	8,573,357	7,138,106	872,859	340,522

Notes: This table contains summary statistics for residents who had their first stays in a nursing home between 2000 and 2010.

Table 2: Summary Statistics for Nursing Homes

Number of Beds	141.3
	(80.72)
Occupancy Percentage	86.44
	(12.01)
Chain	0.611
	(0.487)
For Profit	0.754
	(0.431)
Cited Deficiencies	6.987
	(6.384)
Registered Nurse Staffing (hours per patient-day)	0.365
	(0.389)
Licensed Practical Nurse Staffing (hours per patient-day)	0.830
	(0.535)
Certified Nursing Assistant Staffing (hours per patient-day)	2.225
	(1.087)
Number of Nursing Homes	17,248

Notes: This table contains summary statistics for nursing homes, weighted by the number of residents admitted between 2000 and 2010.

Table 3: Association Between Quality Index and Minority Status

	Quality Index (s.d.)	
	(1)	(2)
Race: Black	-0.120***	-0.110***
	(0.0110)	(0.00487)
Race: Hispanic	-0.140***	-0.0711***
	(0.0168)	(0.00757)
Constant	0.0178**	0.0140**
	(0.00738)	(0.00588)
Zip Code Fixed Effects		X
Number of Observations	8,567,753	8,566,289
R-squared	0.002	0.094

Notes: The unit of observation is a resident. Standard errors are clustered at the nursing home level.

Table 4: Admissions Behavior by Nursing Homes

(a) Evidence of Capacity Constraints

	Number of New Residents		
	(1)	(2)	(3)
Lagged 7-Day Avg. Log Occupancy	-0.298***		
	(0.00527)		
Lagged 7-Day Avg. Occupancy		-0.00935***	
		(0.000107)	
Lagged 7-Day Avg. Occ. Percentile			-0.00246***
			(2.45e-05)
Nursing Home-Month Fixed Effects	X	X	X
N	40,033,450	40,033,450	40,033,450
R-squared	0.205	0.205	0.205

Notes: This table shows regression results at the nursing home-day level wherein the dependent variable is number of new patients, and the independent variables are various measures of nursing home occupancy. Standard errors are clustered at the nursing home level.

(b) Evidence of Selective Admissions (Unconditional)

	Black	Hispanic	Post-Acute
	(1)	(2)	(3)
Lagged 7-Day Avg. Log Occupancy	-0.0256***	-0.00660***	0.0707***
	(0.00290)	(0.00180)	(0.00400)
Nursing Home-Month Fixed Effects	X	X	X
Number of Residents	7,106,929	7,106,929	7,106,929
R-squared	0.385	0.355	0.301

Notes: Regressions are at the resident level. Standard errors are clustered by nursing home.

(c) Evidence of Selective Admissions (Conditional)

	Black	Hispanic	Post-Acute
	(1)	(2)	(3)
Lagged 7-Day Avg. Log Occupancy	-0.0169***	-0.00284	0.0472***
	(0.00285)	(0.00180)	(0.00374)
Nursing Home-Month Fixed Effects	X	X	X
Controls for Other Characteristics	X	X	X
Number of Residents	7,102,426	7,102,426	7,102,426
R-squared	0.393	0.358	0.338

Notes: Regressions are at the resident level, and include controls for race, Medicaid, post-acute care, dementia, age, gender, marital status, and education (as long as the variable is not the dependent variable). Standard errors are clustered by nursing home.

Table 5: Reduced Form Evidence of In-Group Preferences

Panel A: Share of Admitted Residents who are Black			
	OLS	*Anderson-Hsiao*	*Arellano-Bond*
	(1)	(2)	(3)
Black Share (Previous Admits)	0.246***	0.143***	0.0359***
	(0.0133)	(0.0229)	(0.0112)
Nursing Home Fixed Effects	X	Differenced-out	Differenced-out
County x Year Fixed Effects	X	X	X
Number of Nursing Home-Years	114,962	100,608	112,017
Panel B: Share of Admitted Residents who are Hispanic			
	OLS	*Anderson-Hsiao*	*Arellano-Bond*
	(1)	(2)	(3)
Hispanic Share (Previous Admits)	0.184***	0.120***	0.0412**
	(0.0183)	(0.0344)	(0.0172)
Nursing Home Fixed Effects	X	Differenced-out	Differenced-out
County x Year Fixed Effects	X	X	X
Number of Nursing Home-Years	114,962	100,608	112,017

Notes: This table shows regression results at the nursing home-year level, with weights equal to the number of residents who were admitted to the nursing home during that year. The Anderson-Hsiao and Arellano-Bond specifications correspond to dynamic panel methods from Anderson and Hsiao (1982) and Arellano and Bond (1991) respectively. Standard errors are clustered at the nursing home level.

Table 6: Effect of Short-Term Fluctuations in Racial Composition on Admissions of Different Racial Groups

Panel A: White Admissions	(1)	(2)	(3)	(4)
Log of White Occupancy	1.760***			2.215***
	(0.0254)			(0.0486)
Log of Black Occupancy		0.0135***		0.00238
		(0.00195)		(0.00279)
Log of Hispanic Occupancy			0.00922***	0.00446*
			(0.00196)	(0.00230)
Nursing Home-Month Fixed Effects	X	X	X	X
Number of Observations	39,969,813	24,296,281	15,202,405	11,554,850
R-squared	0.298	0.296	0.308	0.325
Panel B: Black Admissions	(1)	(2)	(3)	(4)
Log of White Occupancy	0.00877***			-0.00180
	(0.00159)			(0.00360)
Log of Black Occupancy		0.249***		0.266***
		(0.00254)		(0.00351)
Log of Hispanic Occupancy			0.00135*	0.000580
			(0.000706)	(0.000911)
Nursing Home-Month Fixed Effects	X	X	X	X
Number of Observations	39,969,813	24,296,281	15,202,405	11,554,850
R-squared	0.192	0.189	0.199	0.201
Panel C: Hispanic Admissions	(1)	(2)	(3)	(4)
Log of White Occupancy	0.000191			-0.00774***
	(0.000910)			(0.00276)
Log of Black Occupancy		0.000177		-0.00118*
		(0.000324)		(0.000666)
Log of Hispanic Occupancy			0.204***	0.211***
			(0.00230)	(0.00257)
Nursing Home-Month Fixed Effects	X	X	X	X
Number of Observations	39,969,813	24,296,281	15,202,405	11,554,850
R-squared	0.216	0.216	0.222	0.220

Notes: This table shows regression estimates at the nursing home-day level of admissions of residents of different races on log occupancy of different races, controlling for nursing home-month fixed effects. Standard errors are clustered at the nursing home level.

Table 7: Evidence on Racial Heterogeneity in Sensitivity to Quality

Panel A. Dependent Variable: Indicator for Admitted Resident Being White

	(1)	(2)	(3)	(4)
Quality Index	0.00188***	0.000549	0.00162***	0.000939**
	(0.000317)	(0.000431)	(0.000445)	(0.000383)
White Share (Previous Admits)	0.966***	0.306***	0.246***	0.176***
	(0.00175)	(0.0111)	(0.0107)	(0.00801)
Nursing Home Fixed Effects		X	X	X
County x Year Fixed Effects			X	X
Zip Code of Prior Residence				X
Number of Observations	7,689,908	7,689,788	7,689,739	7,688,164
R-squared	0.293	0.303	0.307	0.397

Panel B. Dependent Variable: Indicator for Admitted Resident Being Black

	(1)	(2)	(3)	(4)
Quality Index	-0.00122***	-0.000825**	-0.00144***	-0.000899***
	(0.000274)	(0.000371)	(0.000374)	(0.000307)
Black Share (Previous Admits)	0.961***	0.278***	0.245***	0.164***
	(0.00203)	(0.0133)	(0.0129)	(0.00871)
Nursing Home Fixed Effects		X	X	X
County x Year Fixed Effects			X	X
Zip Code of Prior Residence				X
Number of Observations	7,689,908	7,689,788	7,689,739	7,688,164
R-squared	0.275	0.286	0.289	0.396

Panel C. Dependent Variable: Indicator for Admitted Resident Being Hispanic

	(1)	(2)	(3)	(4)
Quality Index	-0.000755***	0.000296	-0.000214	-6.02e-05
	(0.000170)	(0.000230)	(0.000251)	(0.000240)
Hispanic Share (Previous Admits)	0.967***	0.235***	0.175***	0.152***
	(0.00413)	(0.0170)	(0.0179)	(0.0170)
Nursing Home Fixed Effects		X	X	X
County x Year Fixed Effects			X	X
Zip Code of Prior Residence				X
Number of Observations	7,689,908	7,689,788	7,689,739	7,688,164
R-squared	0.255	0.266	0.269	0.311

Notes: This table shows regression results at the resident level. Previous share of a given race is defined as the fraction of admissions in the previous year to the nursing home that the resident is admitted to who belong to the race. Standard errors are clustered at the nursing home level.

Table 8: Estimates of Residents' Preferences and Selective Admissions by Nursing Homes

	California	*Florida*	*New York*	*Texas*
Residents' Preferences	(1)	(2)	(3)	(4)
Utility (Distance to Facility)	-0.159***	-0.176***	-0.189***	-0.16***
	(0.001)	(0.001)	(0.001)	(0.001)
Utility (Previous Share Black)	-0.938***	-1.521***	-1.178***	-1.495***
	(0.017)	(0.023)	(0.011)	(0.023)
Utility (Black x Previous Share Black)	1.509***	1.966***	1.613***	1.959***
	(0.028)	(0.045)	(0.021)	(0.039)
Utility (Previous Share Hispanic)	-0.989***	-1.082***	-1.149***	-1.607***
	(0.019)	(0.027)	(0.021)	(0.031)
Utility (Hispanic x Previous Share Hispanic)	2.086***	1.205***	1.839***	1.719***
	(0.027)	(0.038)	(0.032)	(0.058)
Utility (Quality Index)	0.058***	0.115***	0.138***	0.108***
	(0.003)	(0.004)	(0.004)	(0.006)
Utility (Quality Index x Black)	-0.039***	-0.174***	-0.09***	-0.043***
	(0.009)	(0.011)	(0.009)	(0.014)
Utility (Quality Index x Hispanic)	-0.047***	-0.16***	-0.083***	-0.063***
	(0.007)	(0.011)	(0.01)	(0.014)
Nursing Homes' Admission Policies				
Occupancy	-3.137***	-5.024***	-4.904***	-3.504***
	(0.142)	(0.373)	(0.118)	(0.259)
Race (Black)	-1.264***	-0.943***	-0.799***	-1.015***
	(0.049)	(0.063)	(0.033)	(0.051)
Race (Hispanic)	-0.572***	-0.877***	-0.781***	-0.399***
	(0.03)	(0.036)	(0.037)	(0.052)

Notes: This table shows estimates of the structural model using Gibbs sampling. A burn-in period corresponding to the first half of the chain was used. The supply-side equation includes an intercept, the quality index, and a dummy for the year 2010.

B Data Appendix

Data on residents was obtained from the Minimum Data Set 2.0 (MDS), which was then superseded by the Minimum Data Set 3.0 after 2010. I focus on the earlier period given inconsistencies in the variables collected across versions, and the fact that the earlier version contains residents' zip code of prior residence. In principle, zip code for some residents in the later period can be obtained by linking the MDS data with Medicare and Medicaid data, but I will still have to drop residents without Medicare or Medicaid. Data on nursing homes was obtained from the Online Survey Certification and Reporting (OSCAR) data, and was downloaded from LTCFocus, a product of the Shaping Long-Term Care in America Project being conducted at the Brown University Center for Gerontology and Healthcare Research and supported, in part, by the National Institute on Aging. This data contains yearly level information on nursing homes such as the street address, and average RN, LPN, and CNA staffing levels. Additional information on deficiency citations and star ratings were obtained from the CMS website.

In constructing my sample of residents, I drop the relatively small number of residents with errors in birth or death dates (e.g., with different birth or death dates recorded across different assessments). In addition, I drop residents admitted to nursing homes with Medicare Provider Numbers that could not be linked to the OSCAR data. Race for Black and Hispanic individuals is coded based on the corresponding variable in the MDS data, and the base category (which I refer to as White in the main text) includes a small number of Asian and individuals of other races. These races were self-reported by the resident (or their family if the resident has trouble communicating), and only one race category can be chosen. For the structural demand estimation sample, I consider only resident-nursing home pairs within 15 miles of each other. In addition, I drop nursing homes that admit fewer than 30 residents over the period of my structural estimation sample (2008–2010) to ensure sufficient power.

Two key variables for the structural estimation are distance between residents and nursing homes (which serves as the demand-side instrument), and temporary occupancy fluctuations (which I use as the supply-side instrument). To construct distances, I combine residents' zip code from the MDS with nursing homes' street address from the OSCAR data, which I convert to latitude and longitude using the Google Maps API. The stata command "geodist" was then used to compute the distance between residents' zip code of prior residence, and nursing homes' locations.

For temporary occupancy fluctuations, I consider the average log occupancy (or average occupancy or within-nursing home occupancy percentile as robustness checks) of nursing homes within 15 miles of each resident in the week before the resident was admitted to the nursing home. I focus on the week prior to admissions instead of a shorter term measure such as the day of/before admission given that residents (or hospital discharge planners) typically need some time to search for and coordinate with nursing homes. In addition, I residualize lagged 7-day log occupancy of nursing home-month fixed effects, to abstract from expansions or contractions that a nursing home may be undergoing (for example, if a nursing home is expanding, it will be more likely to admit new residents even though its occupancy seems high). The OSCAR contains data on total number of beds at a nursing home which one could in principle use as a measure of capacity, but this variable is measured with substantial error and only updated annually. In fact, measures of occupancy over time based on residents' admission and discharge dates from the MDS frequently exceeds total number of beds reported by nursing homes in the OSCAR data, which calls into question the use of this variable as a measure of capacity.

C Dynamic Panel Methods

Recall the structural equation in equation (2) studying the relationship between past shares of admissions minority and present share of admissions minority:

$$s_{jt}^{r} = \alpha^{r,ingroup} + \beta^{r,ingroup} s_{j,t-1}^{r} + \delta_{j}^{r,ingroup} + \gamma_{ct}^{r,ingroup} + \epsilon_{jt}^{r,ingroup}.$$

A concern with OLS estimation of this equation is that since this is a relatively short panel, the inclusion of nursing home fixed effects may lead to the Nickell (1981) bias if demeaning is used to estimate the model coefficients. Hence, I also estimate specifications based on the GMM estimators proposed in Anderson and Hsiao (1982) and Arellano and Bond (1991).

To describe these methods, consider the first-differenced structural equation:

$$\Delta s_{jt}^{r} = \beta^{r,ingroup} \Delta s_{j,t-1} + \Delta \gamma_{ct}^{r,ingroup} + \Delta \epsilon_{jt}^{r,ingroup},$$

where $\Delta U_{jt} \equiv U_{jt} - U_{j,t-1}$ for any random variable U. Under the weak exogeneity condition, we have:

$$\epsilon_{jt}^{r,ingroup} \perp \left(s_{j}^{r,(t-2)\prime}, \delta_{j}^{r,ingroup\prime}, \gamma_{c}^{r,ingroup,(t-1)\prime} \right)',$$

where $U_{j}^{(t)} \equiv (U_{jt}, U_{j,t-1}, ...)$. This generates moment conditions that we can use to estimate equation

2:

$$\mathbb{E}\left[\left(\Delta s^{r}_{jt}-\beta^{r,ingroup}\Delta s_{j,t-1}-\Delta\gamma^{r,ingroup}_{ct}\right)\left(s^{(t-2)\prime}_{j},\Delta\gamma^{r,ingroup,(t-1)\prime}_{c}\right)'\right]=0.$$

The Anderson-Hsiao estimation approach only uses the moment condition generated by the first lag of the pre-determined regressors, i.e.,

$$\mathbb{E}\left[\left(\Delta s^{r}_{jt}-\beta^{r,ingroup}\Delta s_{j,t-1}-\Delta\gamma^{r,ingroup}_{ct}\right)\left(s_{j,t-2},\Delta\gamma^{r,ingroup}_{c,t-1}\right)'\right]=0,$$

whereas the Arellano-Bond approach uses all available moment conditions.

D Algorithm for Gibbs Sampler

In the following description for the Gibbs sampler, when drawing structural error terms in sequence for $j\in\mathcal{J}_i$, I assume an increasing order (although obviously any other order works as well). In addition, to simplify notation, I denote variables in residents' utility and nursing homes' admissions equations by $\mathbf{X}_{ij}$ and W_{ij} respectively,[17] and refer to the nursing home that resident i ends up in by $\mu(i)$.

Denoting iterations of the Gibbs sampler by k and indicating the values of various parameters in the kth iteration of the Gibbs sampler using a superscript k, the steps for implementing the Gibbs sampler are as follows.

1. Initialization ($k=0$): I assume that $(\epsilon_{ij},\omega_{ij})\sim^{i.i.d.}N(0,I_2)$ and set the following conjugate priors for the parameters: $(\kappa',\psi')'\sim N(0,100I)$.
   (a) Set the initial values of the parameters $\theta^0=\left(\kappa^{0\prime},\psi^{0\prime}\right)$ at their prior mean.
   (b) Initial data augmentation: For each resident i, draw the vector ϵ^0_i such that $v^0_{i,\mu(i)}\geq v^0_{ij}$ for all $j\in\mathcal{J}_i$.
      i. Draw $\omega^0_{i,\mu(i)}$ such that $\omega^0_{i,\mu(i)}\geq -W'_{ij}\psi^0$ and for $j\neq\mu(i)$ draw ω^0_{ij} from the unconditional distribution.
      ii. Set $\epsilon^0_{i,\mu(i)}$ equal to three times the standard deviation of the prior. For $j\neq\mu(i)$, draw ϵ^0_{ij} such that $\epsilon^0_{ij}\leq\left(\mathbf{X}_{i,\mu(i)}-\mathbf{X}_{ij}\right)'\kappa^0+\epsilon^0_{i,\mu(i)}$ if $\pi^0_{ij}\geq 0$ or draw ϵ^0_{ij} unconditionally otherwise.
2. For $k+1=1,...,K$:

[17] These include resident characteristics x_i, nursing home characteristics w_j, distance between residents and nursing homes $dist_{ij}$, occupancy fluctuations at nursing homes occ_{ij}, and interactions between these variables.

(a) Draw the profit shocks $\omega_i^{k+1}|v_i^k;\psi^k$ in sequence for $j \in \mathcal{J}_i$.

i. If $v_{ij}^k < v_{i,\mu(i)}^k$, draw ω_{ij}^{k+1} unconditional on assignment (given that even if i is eligible for j, i would not choose j).

ii. If $v_{ij}^k > v_{i,\mu(i)}^k$, draw ω_{ij}^{k+1} from a truncated normal with mean and variance given by the conditional distribution and truncation point $\omega_{ij}^{k+1} < -W_{ij}\psi^k$ (given that otherwise i would choose j over $\mu(i)$).

iii. Finally, if $j = \mu(i)$, draw from the conditional distribution with truncation point given by $\omega_{ij}^{k+1} \geq -W_{ij}'\psi^k$ (given that i must always be eligible for the facility she was ultimately assigned to).

(b) Update π_i^{k+1} according to $\pi_{ij}^{k+1} = W_{ij}'\psi^k + \omega_{ij}^{k+1}$.

(c) Draw the utility shocks $\epsilon_i^{k+1}|\pi_i^{k+1};\kappa^k$ in sequence, for $j \in \mathcal{J}_i$.

i. If $\pi_{ij}^{k+1} < 0$, draw ϵ_{ij}^{k+1} unconditionally (given that i would not choose such a facility even if she were eligible for it).

ii. If $\pi_{ij}^{k+1} \geq 0$ and $j \neq \mu(i)$, draw ϵ_{ij}^{k+1} from the conditional distribution with truncation point given by $v_{ij}^{k+1} < \mathbf{X}_{ij}'\kappa^k$.

iii. For $j = \mu(i)$, draw $\epsilon_{i,\mu(i)}^{k+1}$ such that $v_{i,\mu(i)}^{k+1}$ is larger than the current values of $v_{i,j'}$ for $j' \neq j$ and $\pi_{ij'} \geq 0$.

(d) Update v_i^{k+1} according to $v_{ij}^{k+1} = \mathbf{X}_{ij}'\kappa^k + \epsilon_{ij}^{k+1}$.

(e) Update the parameters θ based on the new indirect utilities v^{k+1} and profits π^{k+1}.

i. First, we update κ. Denote the design matrix in the equation for indirect utilities by $\mathbb{X}$. In matrix notation, we have:

$$v = \mathbb{X}\kappa + \epsilon,\ \epsilon \sim N(0, I).$$

We have a normal conjugate prior for κ, with mean μ_κ^0 and covariance matrix Σ_κ^0. The posterior distribution of κ conditional on v and W is:

$$\kappa|(v, \mathbb{X}) \sim N(\tilde{\mu}_\kappa, \tilde{\Sigma}_\kappa),$$

where the posterior mean and covariance matrix are given by:

$$
\begin{aligned}
\tilde{\mu}_\kappa &= \left(\frac{\mathbb{X}'\mathbb{X}}{\sigma_\epsilon^2} + \left(\Sigma_\kappa^0\right)^{-1}\right)^{-1} \left(\left(\Sigma_\kappa^0\right)^{-1}\mu_\kappa^0 + \frac{\mathbb{X}'\kappa}{\sigma_\epsilon^2}\right) \\
&= \left(\mathbb{X}'\mathbb{X} + \left(\Sigma_\kappa^0\right)^{-1}\right)^{-1} \left(\left(\Sigma_\kappa^0\right)^{-1}\mu_\kappa^0 + \mathbb{X}'\kappa\right),
\end{aligned}
$$

$$
\begin{aligned}
\tilde{\Sigma}_{\theta_v} &= \left(\frac{\mathbb{X}'\mathbb{X}}{\sigma_\epsilon^2} + \left(\Sigma_\kappa^0\right)^{-1}\right)^{-1} \\
&= \left(\mathbb{X}'\mathbb{X} + \left(\Sigma_\kappa^0\right)^{-1}\right)^{-1}.
\end{aligned}
$$

We then set κ^{k+1} by drawing from this posterior distribution.

A. Next, we will update ψ. Denote the design matrix in the equation for the admissions rule by W. In matrix notation, we have:

$$\pi = W\psi + \omega,\ \omega \sim N(0, I).$$

We have a normal prior for ψ, with mean μ_ψ^0 and covariance matrix Σ_ψ^0, so the posterior distribution of θ_π conditional on π and W is:

$$\psi|(\pi, W) \sim N(\tilde{\mu}_\psi, \tilde{\Sigma}_\psi),$$

with posterior mean and covariance matrices given by:

$$
\begin{aligned}
\tilde{\mu}_\psi &= \left(\frac{W'W}{\sigma_\omega^2} + \left(\Sigma_\psi^0\right)^{-1}\right)^{-1} \left(\left(\Sigma_\psi^0\right)^{-1}\mu_\psi^0 + \frac{W'\psi}{\sigma_\omega^2}\right) \\
&= \left(W'W + \left(\Sigma_\psi^0\right)^{-1}\right)^{-1} \left(\left(\Sigma_\psi^0\right)^{-1}\mu_\psi^0 + W'\psi\right),
\end{aligned}
$$

$$
\begin{aligned}
\tilde{\Sigma}_\psi &= \left(\frac{W'W}{\sigma_\omega^2} + \left(\Sigma_\psi^0\right)^{-1}\right)^{-1} \\
&= \left(W'W + \left(\Sigma_\psi^0\right)^{-1}\right)^{-1}.
\end{aligned}
$$

We then set ψ^{k+1} by drawing from this posterior distribution.

E Simulation Details

Recall that our structural model is based on equations for residents' decision utility and nursing homes' admission rules respectively:

$$\begin{aligned} v_{ij} =& \kappa_0^{black} s_{ij}^{black} + \kappa_0^{hisp} s_{ij}^{hisp} + \kappa_1^{black} s_{ij}^{black} black_i + \kappa_1^{hisp} s_{ij}^{hisp} hispanic_i \\ & + q_j' \kappa_0^q + black_i q_j' \kappa_{black}^q + hispanic_i q_j' \kappa_{hisp}^q + dist_{ij}' \kappa^{dist} + \epsilon_{ij}, \end{aligned}$$

$$\pi_{ij} = \psi_0 + \psi^{black} black_i + \psi^{hisp} hispanic_i + \tilde{x}_i' \psi^{\tilde{x}} + occ_{ij}' \psi^{occ} + \omega_{ij}.$$

Nursing home j is willing to admit resident i if and only if $\pi_{ij} \geq 0$, and resident i chooses the nursing home which yields the highest decision utility among the set of nursing homes that are willing to admit her. For computational feasibility, for each resident i, I only consider nursing homes within 15 miles of her $\mathcal{J}_i \equiv \{j | dist_{ij} \leq 15\ miles\}$. I denote estimated using "hats", e.g., $(\hat{\kappa}', \hat{\psi}')'$, but for the counterfactuals I switch to using "stars".

To simulate the elimination of in-group preferences, I set $\kappa_0^{r*} = \kappa_1^{r*} = 0$, and to simulate the elimination of racial heterogeneity in sensitivity to quality, I set $\kappa_r^{q*} = 0$. Similarly, to simulate the elimination of discriminatory admissions practices, I set $\psi^{r*} = 0$. To simulate the elimination of residential segregation, I permute the zip codes of prior address for residents. Hence, counterfactual distances between resident i and different nursing homes $dist_{ij}^*$ will generally differ from the original distances $dist_{ij}$, and i is faced with a different potential choice set $\mathcal{J}_i^* \equiv \{j | dist_{ij}^* \leq 15\ miles\}$, unless she is randomized to the same zip code. By virtue of the permutation process, the unconditional geographical distribution of residents' prior addresses remains unchanged, but the distribution of races within each zip code will reflect the overall distribution of race (in expectation), hence eliminating residential segregation.

In terms of notation, I will use κ^*, ψ^*, and $dist^*$ throughout the description of simulations, and it should be understood that this is either equal to its estimated or original value if the corresponding component of the simulation is not turned on, or equal to a counterfactual value otherwise. For example, if we are considering a counterfactual with no residential segregation, $dist_{ij}^*$ will generally be different from $dist_{ij}$, whereas in a counterfactual where we take residential segregation as given, I still use the same notation $dist_{ij}^*$, but this will be equal to $dist_{ij}$.

The simulation algorithm is as follows:

1. Setup:

   (a) For each nursing home j, I set the share of j's admissions that are of race r in the last 365 days prior to the start of the simulation to the mean of this value over the estimation period.

   (b) For each nursing home j, I assume that its quality measures are time-invariant over the simulation period, and set this equal to the average over the estimation period.

   (c) For the simulations, I set the number of new arrivals each day to the mean in the data, which is $N_d^* = 177$.

2. Simulation: for day $d^* = 1, ..., D^* = 5000$ of the simulation:

   (a) If the counterfactual assumes no residential segregation, I permute the zip codes of prior address for residents.

   (b) I then randomly select $N_{d^*}^*$ residents and simulate their choices. For resident $i^* = 1, ..., N_d^*$:

      i. I draw $\epsilon_{ij} \sim N(0, 1)$ and compute:

$$\begin{aligned} v_{i^*j} =& \kappa_0^{black*} s_{d^*j}^{black*} + \kappa_0^{hisp*} s_{d^*j}^{hisp} + \kappa_1^{black*} s_{d^*j}^{black} black_i + \kappa_1^{hisp} s_{d^*j}^{hisp} hispanic_i \\ &+ q_j' \kappa_0^{q*} + black_i q_j' \kappa_{black}^{q*} + hispanic_i q_j' \kappa_{hisp}^{q*} + dist_{ij}'^{*} \kappa^{dist*} + \epsilon_{ij}, \end{aligned}$$

      for each nursing home $j \in \mathcal{J}_{i^*}^* \equiv \{j' | dist_{ij'}^* \leq 15 \ miles\}$.

      ii. Also, for each $j \in \mathcal{J}_{i^*}^*$, I draw $\omega_{i^*j} \sim N(0, 1)$ and compute:

$$\pi_{i^*j} = \psi_0^* + \psi^{black*} black_{i^*} + \psi^{hisp*} hispanic_{i^*} + \tilde{x}_{i^*}' \psi^{\tilde{x}*} + occ_{d(i^*)j}'^{*} \psi^{occ*} + \omega_{i^*j}.$$

      iii. I set i^*'s nursing home to be $\mu(i^*) \equiv argmax_j \{v_{i^*j} | \pi_{i^*j} \geq 0, j \in \mathcal{J}_{i^*}^*\}$.[18]

   (c) Next, I update the shares of residents admitted to each nursing home that is of each race in the 365 days leading up to the next day:

      i. Letting N_{dj}^r be the number of residents of race r that is admitted to nursing home j on day d of the simulation and N_{dj} be the number of residents of any race that is admitted to j on day d, I set:

$$s_{d^*+1,j}^{r*} \equiv \frac{\sum_{d=d^*-364}^{d^*} N_{dj}^r}{\sum_{d=d^*-364}^{d^*} N_{dj}}.$$

[18] If $\pi_{i^*j} < 0$ for all $j \in \mathcal{J}^*{}_{i^*}$, then I simply drop the resident, but this occurs extremely rarely in any of the simulations.

(d) Finally, I update the occupancy measures of each nursing home for the next day in the simulation:

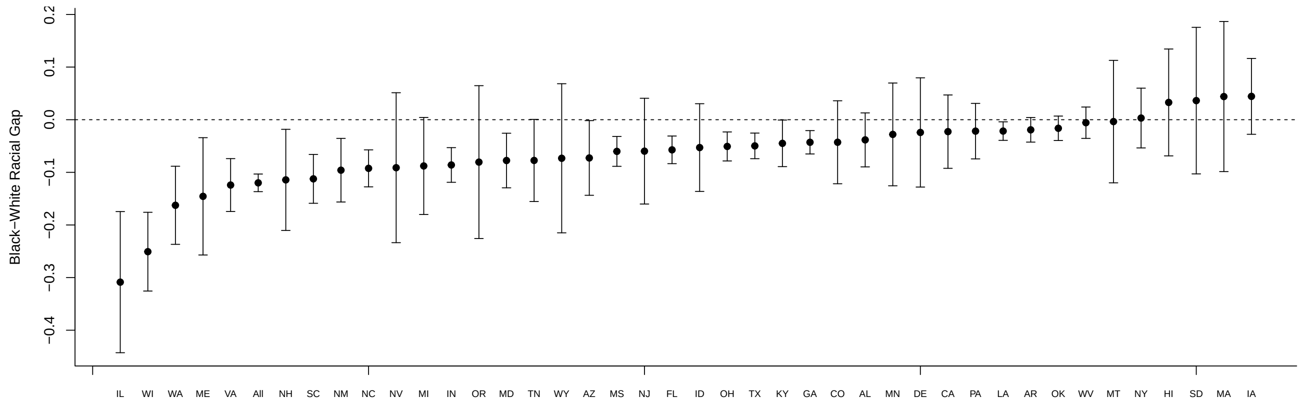
$$occ^*_{d^*+1,j} = log\left(\sum_{d=d^*-6}^{d^*} N_{dj}\right) - log(\bar{N}_j),$$

where $\bar{N}_j$ is the mean occupancy measure at nursing home j in the data.

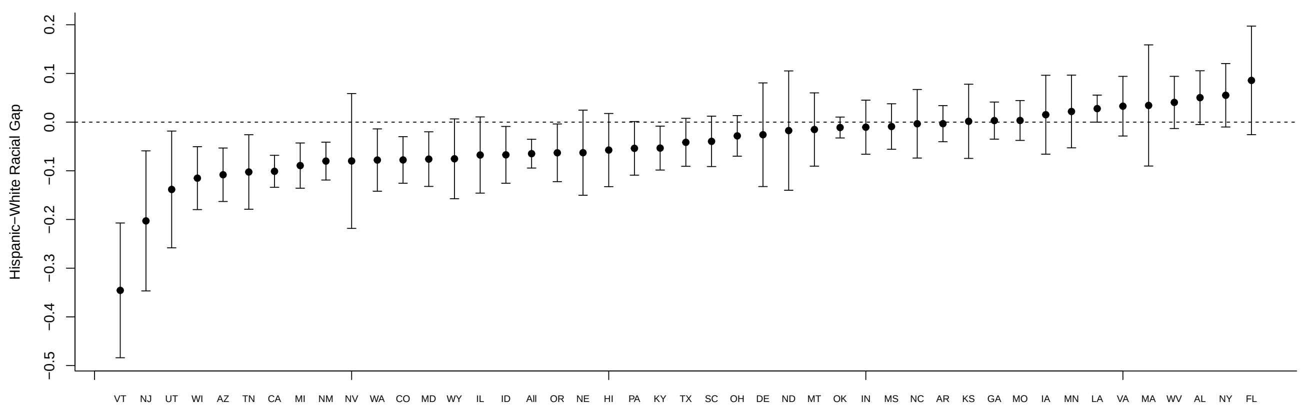
3. Measuring segregation and disparities: to measure segregation and disparities on a day d^* for the simulations, I use data from the past 100 days (i.e., days $d^* - 99$ up to day d^*).

A Appendix Figures and Tables

Figure A.1: Racial Gaps in Nursing Home Quality as Measured by RN Staffing (Standardized)



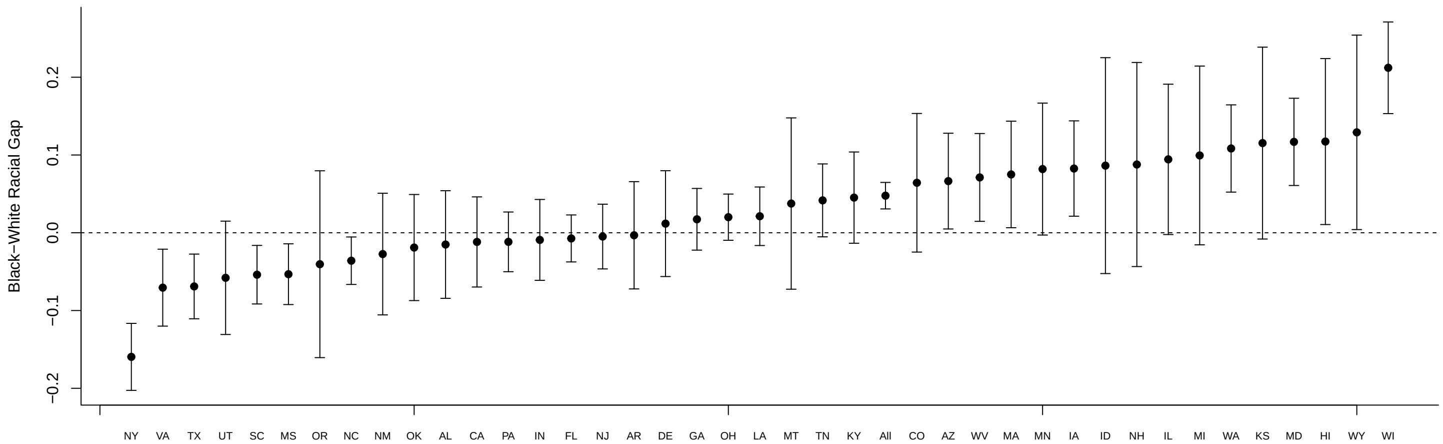
(a) Black-White Gap



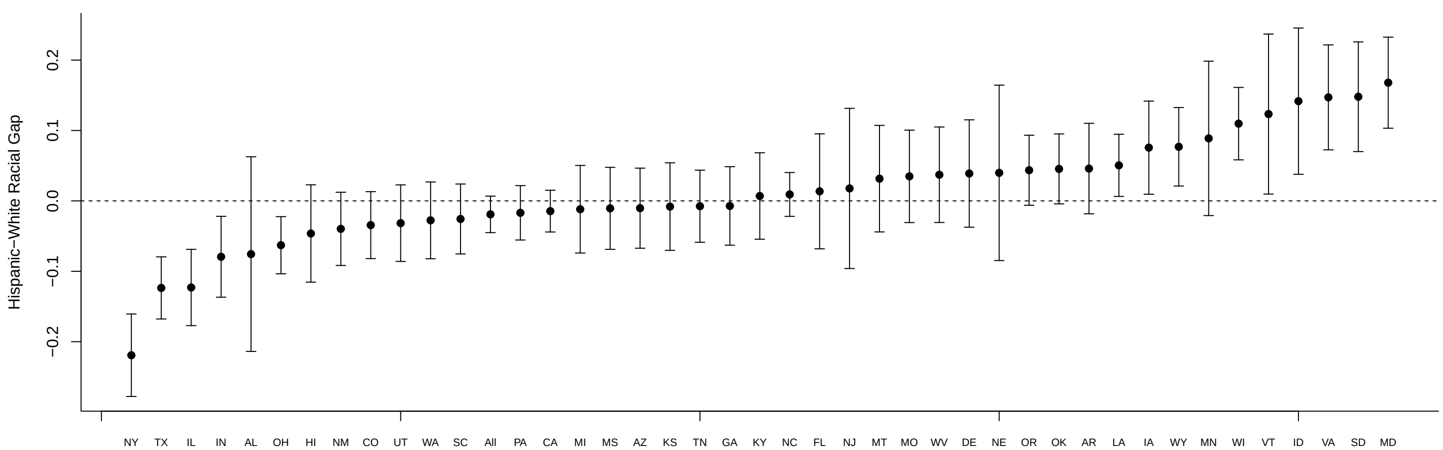
(b) Hispanic-White Gap

Notes: These figures display the estimated racial gaps in nursing home quality by state when RN staffing (standardized to have zero mean and unit variance) is used as the quality measure. Error bars indicate 95 percent confidence intervals for the estimates.

Figure A.2: Racial Gaps in Nursing Home Quality as Measured by LPN Staffing (Standardized)



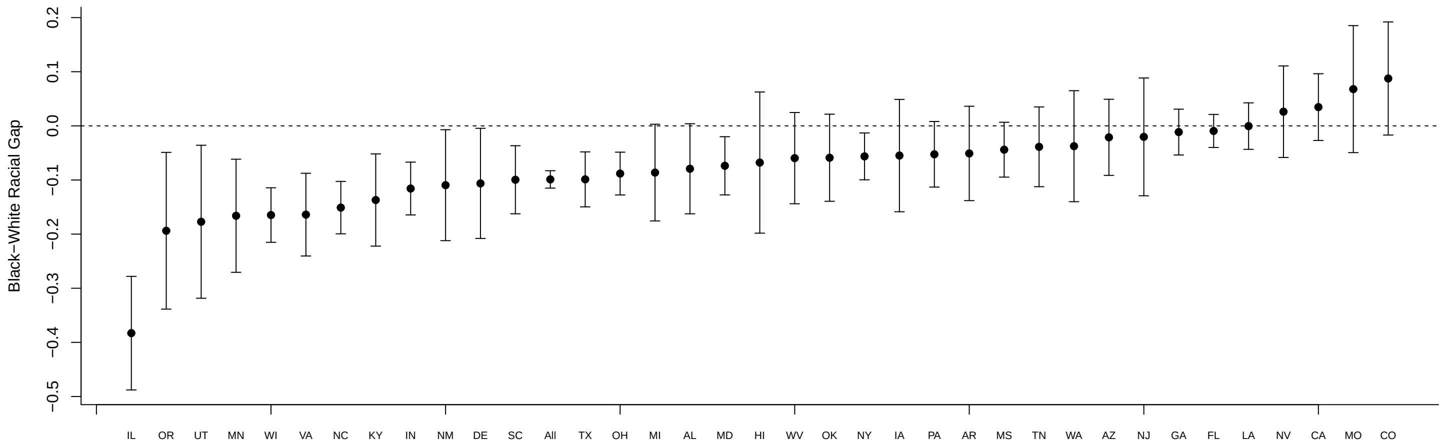
(a) Black-White Gap



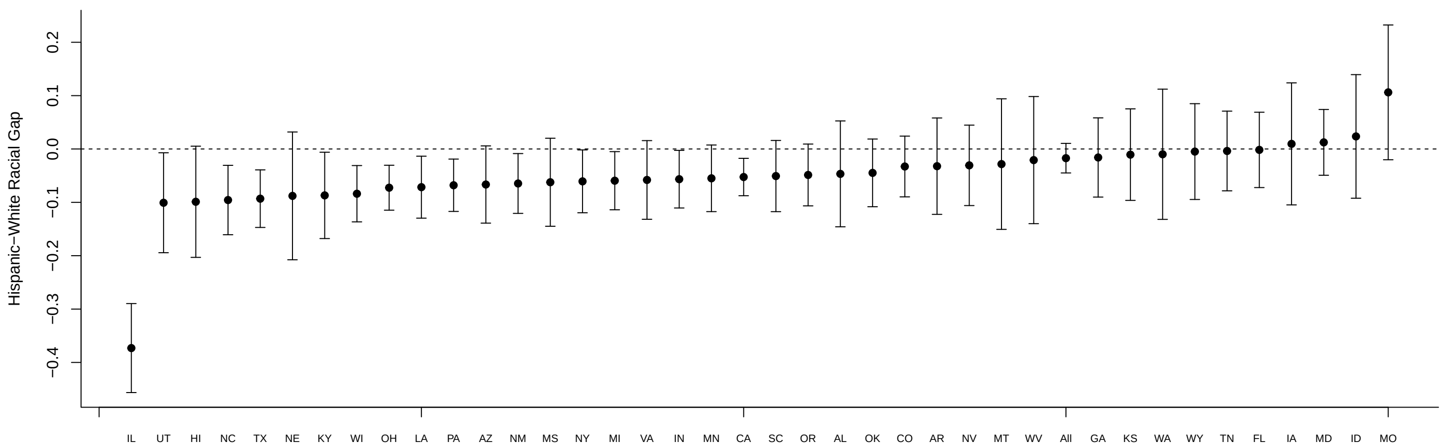
(b) Hispanic-White Gap

Notes: These figures display the estimated racial gaps in nursing home quality by state when LPN staffing (standardized to have zero mean and unit variance) is used as the quality measure. Error bars indicate 95 percent confidence intervals for the estimates.

Figure A.3: Racial Gaps in Nursing Home Quality as Measured by CNA Staffing (Standardized)



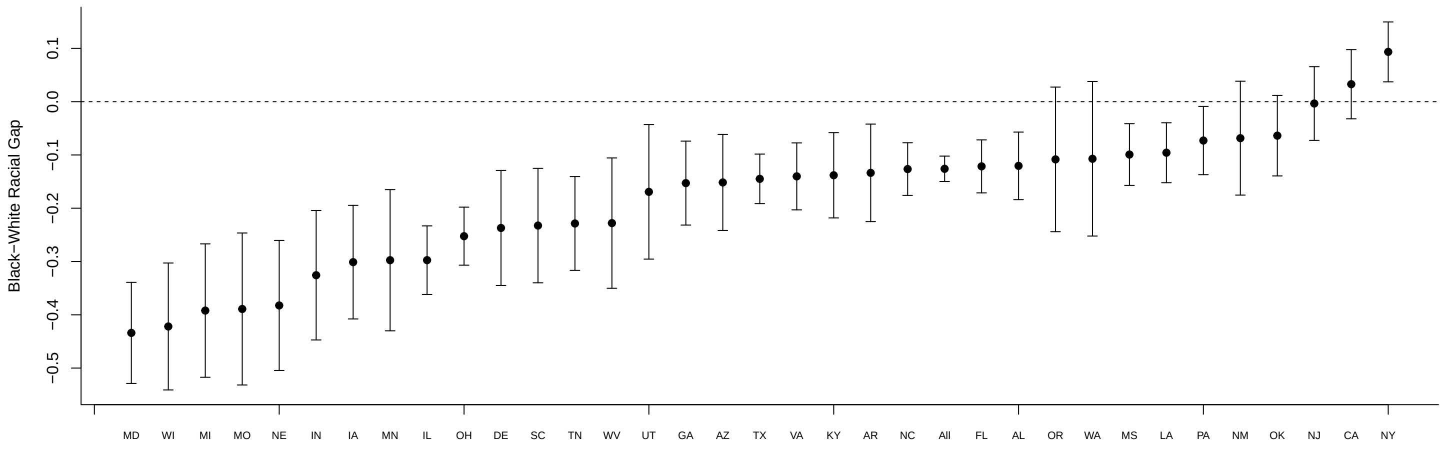
(a) Black-White Gap



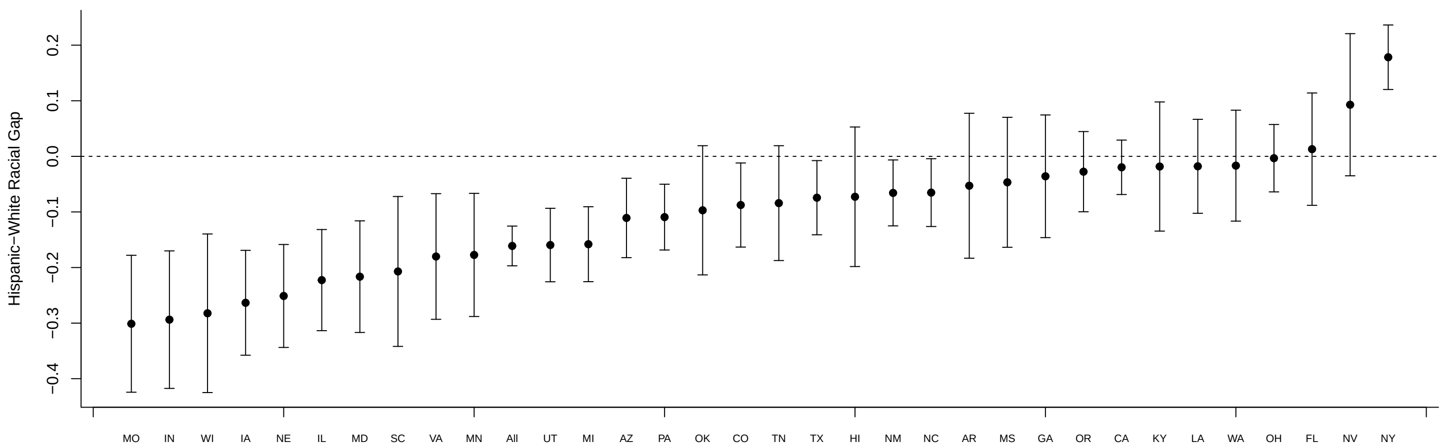
(b) Hispanic-White Gap

Notes: These figures display the estimated racial gaps in nursing home quality by state when CNA staffing (standardized to have zero mean and unit variance) is used as the quality measure. Error bars indicate 95 percent confidence intervals for the estimates.

Figure A.4: Racial Gaps in Nursing Home Quality as Measured by Fewer Standard Deficiencies (Standardized)



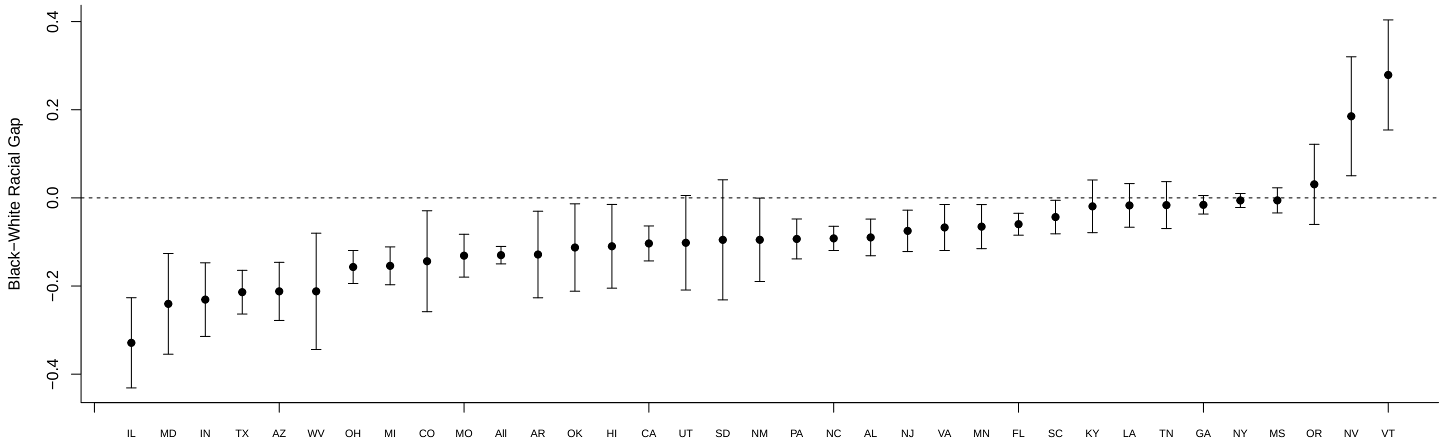
(a) Black-White Gap



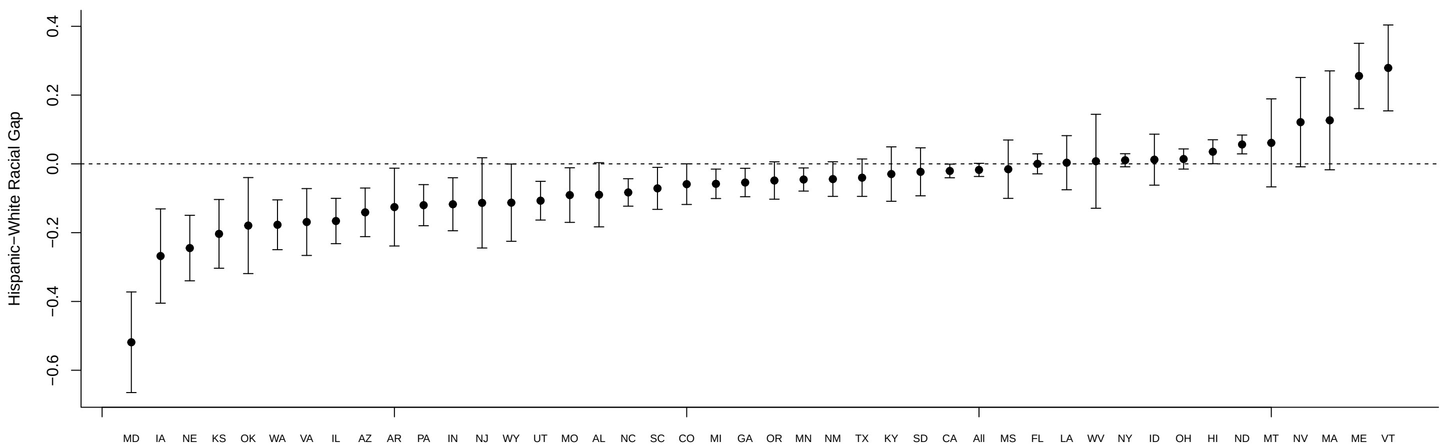
(b) Hispanic-White Gap

Notes: These figures display the estimated racial gaps in nursing home quality by state when the negative of standard deficiencies (standardized to have zero mean and unit variance) is used as the quality measure. Error bars indicate 95 percent confidence intervals for the estimates.

Figure A.5: Racial Gaps in Nursing Home Quality as Measured by Fewer Complaint Deficiencies (Standardized)



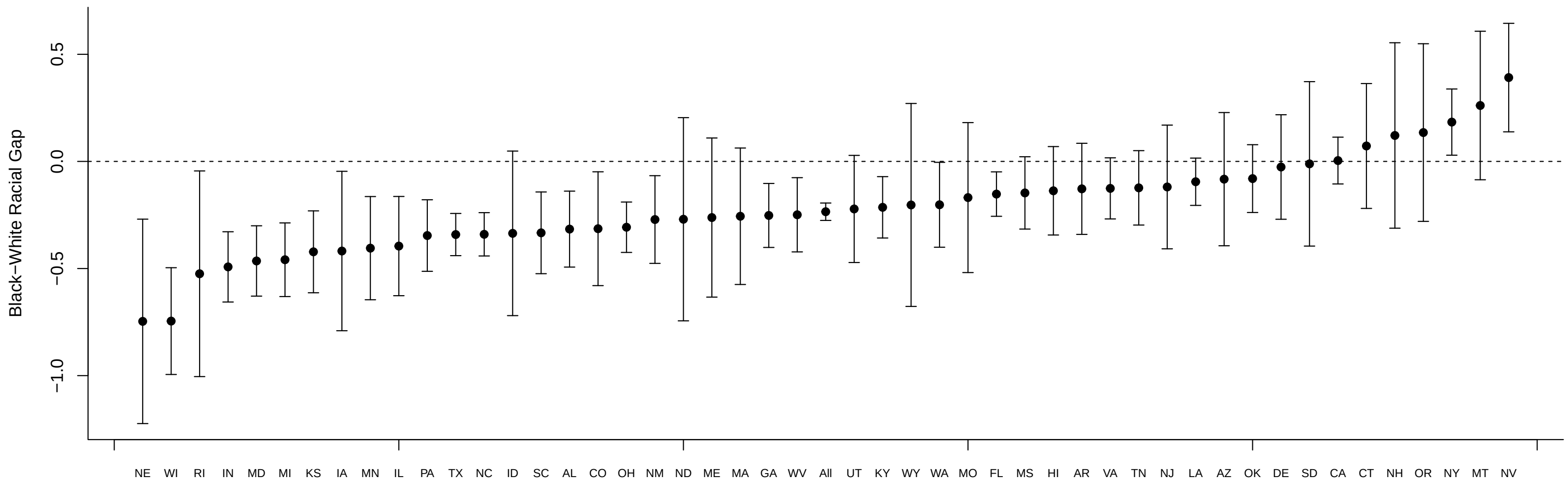
(a) Black-White Gap



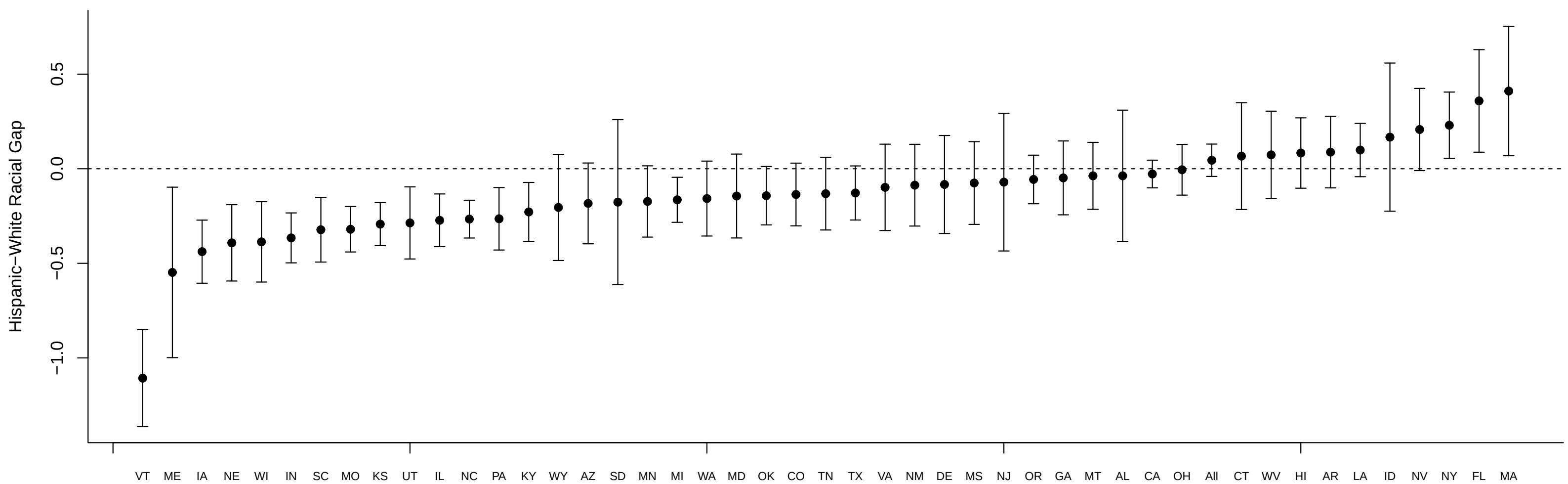
(b) Hispanic-White Gap

Notes: These figures display the estimated racial gaps in nursing home quality by state when the negative of complaint deficiencies (standardized to have zero mean and unit variance) is used as the quality measure. Error bars indicate 95 percent confidence intervals for the estimates.

Figure A.6: Racial Gaps in Nursing Home Quality as Measured by Star Ratings (Standardized)



(a) Black-White Gap

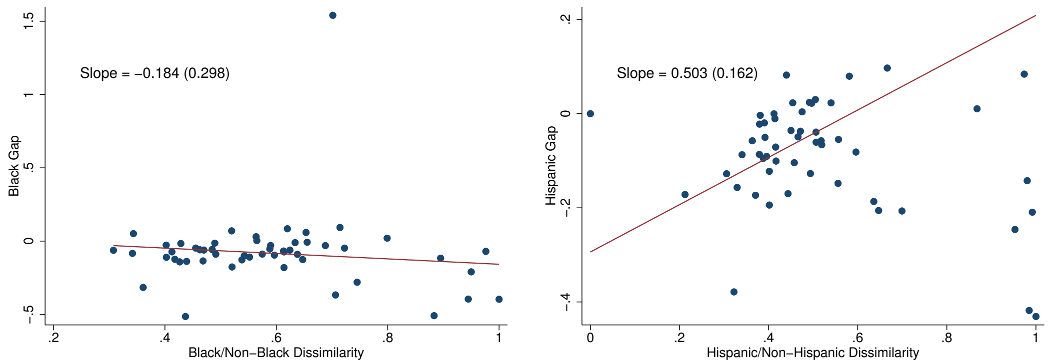


(b) Hispanic-White Gap

Notes: These figures display the estimated racial gaps in nursing home quality by state when 2009 star ratings (standardized to have zero mean and unit variance) is used as the quality measure. Error bars indicate 95 percent confidence intervals for the estimates.

Figure A.7: Cross-Sectional Relationship Between State-Level Segregation and Disparities (RN Staffing)

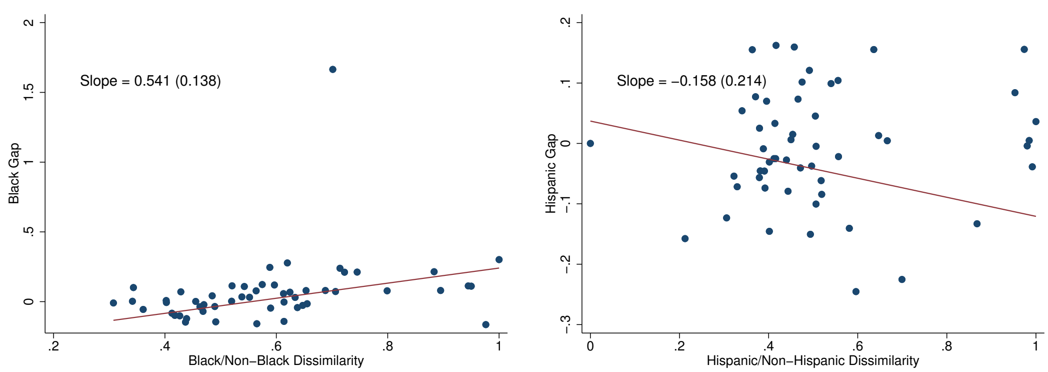
(a) Black/Non-Black (b) Hispanic/Non-Hispanic



Notes: These figures display scatter plots of the estimated racial gap (based on standardized RN staffing) against segregation at the state level. Observations are weighted by the number of residents admitted to the state.

Figure A.8: Cross-Sectional Relationship Between State-Level Segregation and Disparities (LPN Staffing)

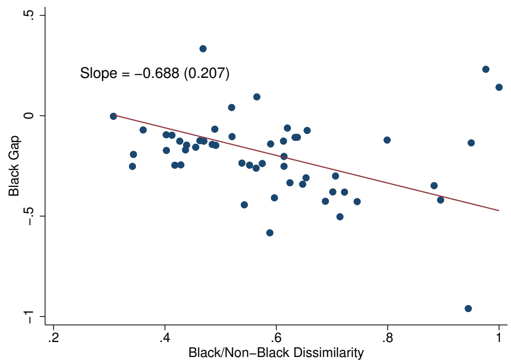
(a) Black/Non-Black (b) Hispanic/Non-Hispanic



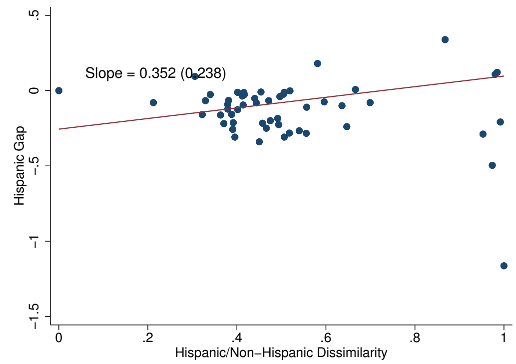
Notes: These figures display scatter plots of the estimated racial gap (based on standardized LPN staffing) against segregation at the state level. Observations are weighted by the number of residents admitted to the state.

Figure A.9: Cross-Sectional Relationship Between State-Level Segregation and Disparities (Fewer Standard Deficiencies)

(a) Black/Non-Black



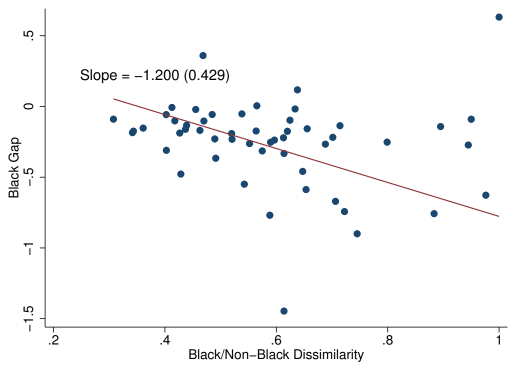
(b) Hispanic/Non-Hispanic



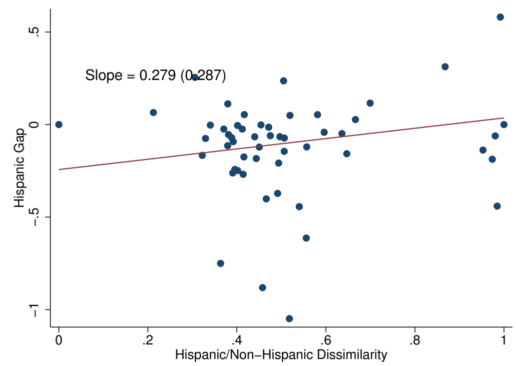
Notes: These figures display scatter plots of the estimated racial gap (based on standardized fewer standard deficiencies) against segregation at the state level. Observations are weighted by the number of residents admitted to the state.

Figure A.10: Cross-Sectional Relationship Between State-Level Segregation and Disparities (Fewer Complaint Deficiencies)

(a) Black/Non-Black

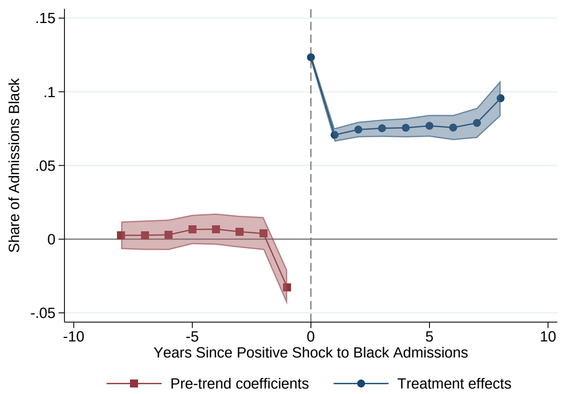


(b) Hispanic/Non-Hispanic

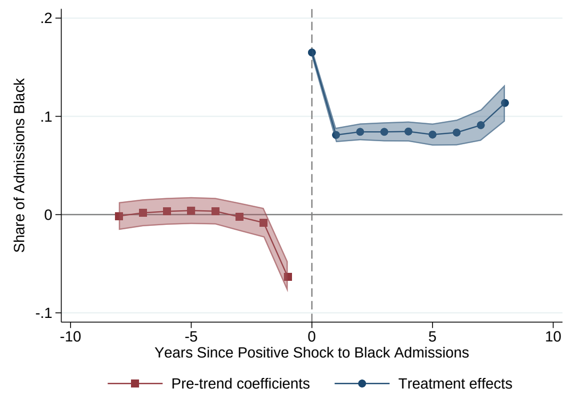


Notes: These figures display scatter plots of the estimated racial gap (based on standardized fewer complaint deficiencies) against segregation at the state level. Observations are weighted by the number of residents admitted to the state.

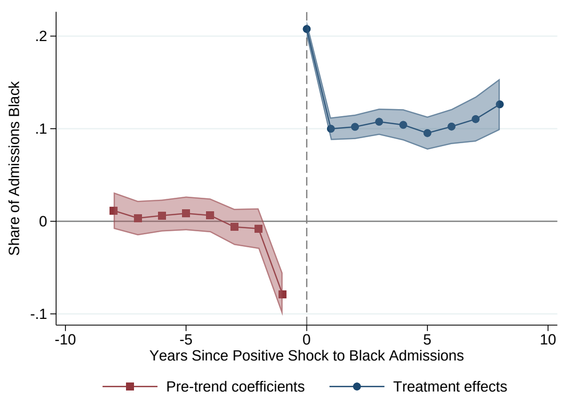
Figure A.11: Event Study on the Effect of a Positive Shock to the Share of Minority Admissions (Other Event Thresholds)



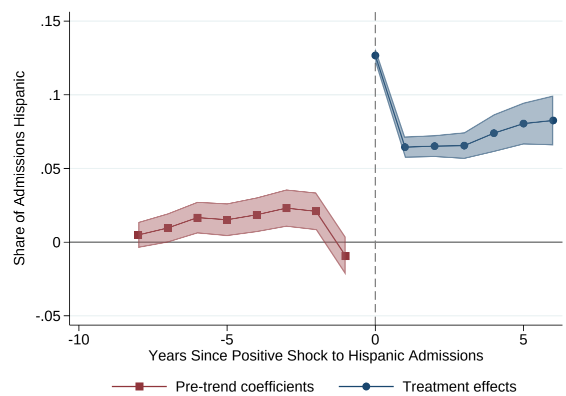
(a) >10pp. Increase in Share of Black Admissions



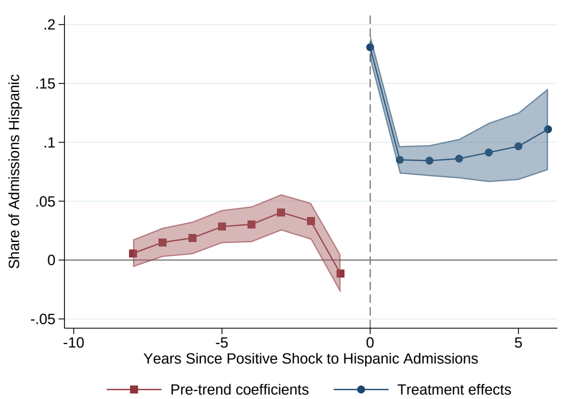
(b) >15pp. Increase in Share of Black Admissions



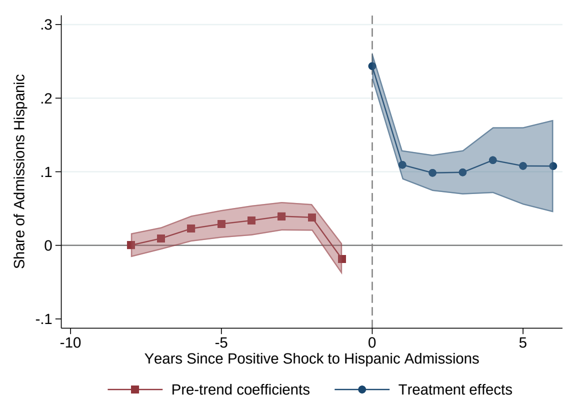
(c) >20pp. Increase in Share of Black Admissions



(d) >10pp. Increase in Share of Hispanic Admissions



(e) >15pp. Increase in Share of Hispanic Admissions

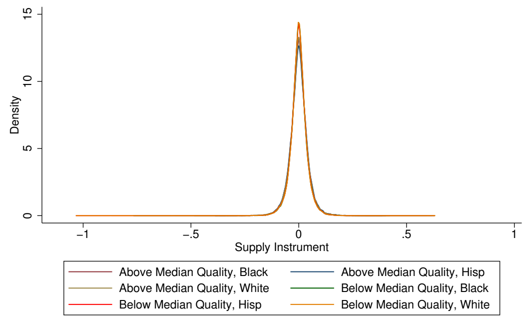


(f) >20pp. Increase in Share of Hispanic Admissions

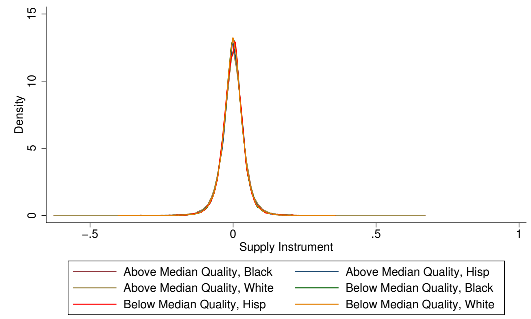
Notes: These figures show event study estimates of the effect of a large shock to the minority share of admissions (defined as a year-to-year increase in the minority share of admissions of at least 10, 15, or 20pp.) on future minority share of admissions, based on event study methods of Borusyak, Jaravel, and Spiess (2021).

Figure A.12: Exclusion Restriction for Temporary Occupancy Fluctuations (Split by RN Staffing and Race of Resident)

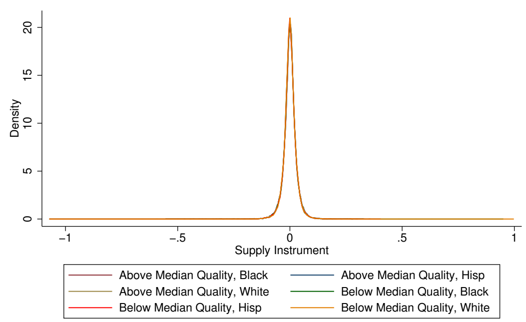
(a) California



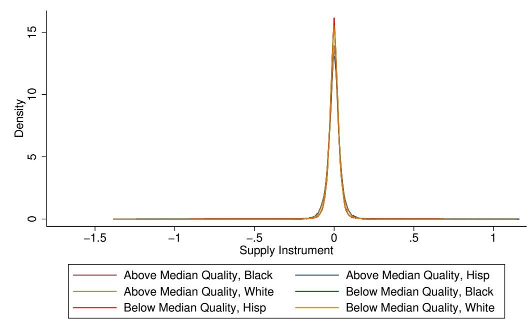
(b) Florida



(c) New York



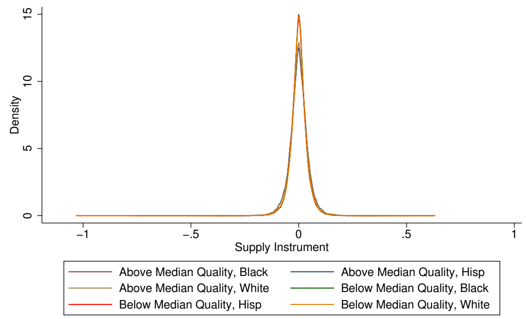
(d) Texas



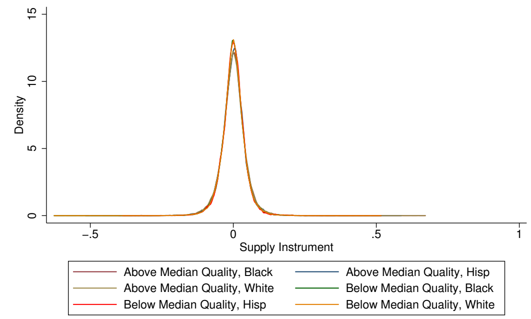
Notes: This figure display kernel density plots of temporary occupancy fluctuations in the four states used for the structural estimation (defined as lagged 7-day log occupancy residualized of nursing home-month fixed effects) of nursing homes within 15 miles of each resident at their time of admission. The unit of observation is a nursing home-resident pair, and the sample is split by whether nursing homes are above-median or below-median in quality (based on RN staffing), and the race of the resident.

Figure A.13: Exclusion Restriction for Temporary Occupancy Fluctuations (Split by LPN Staffing and Race of Resident)

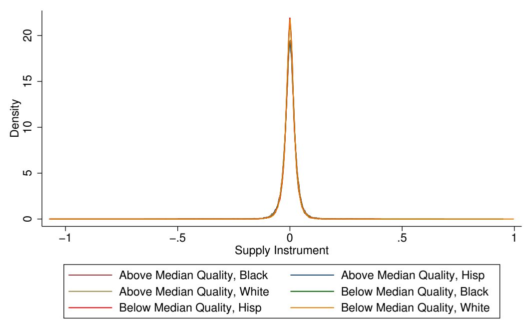
(a) California



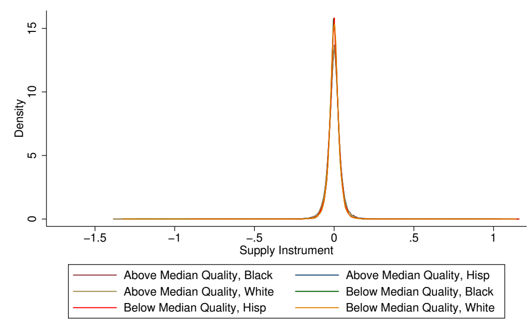
(b) Florida



(c) New York

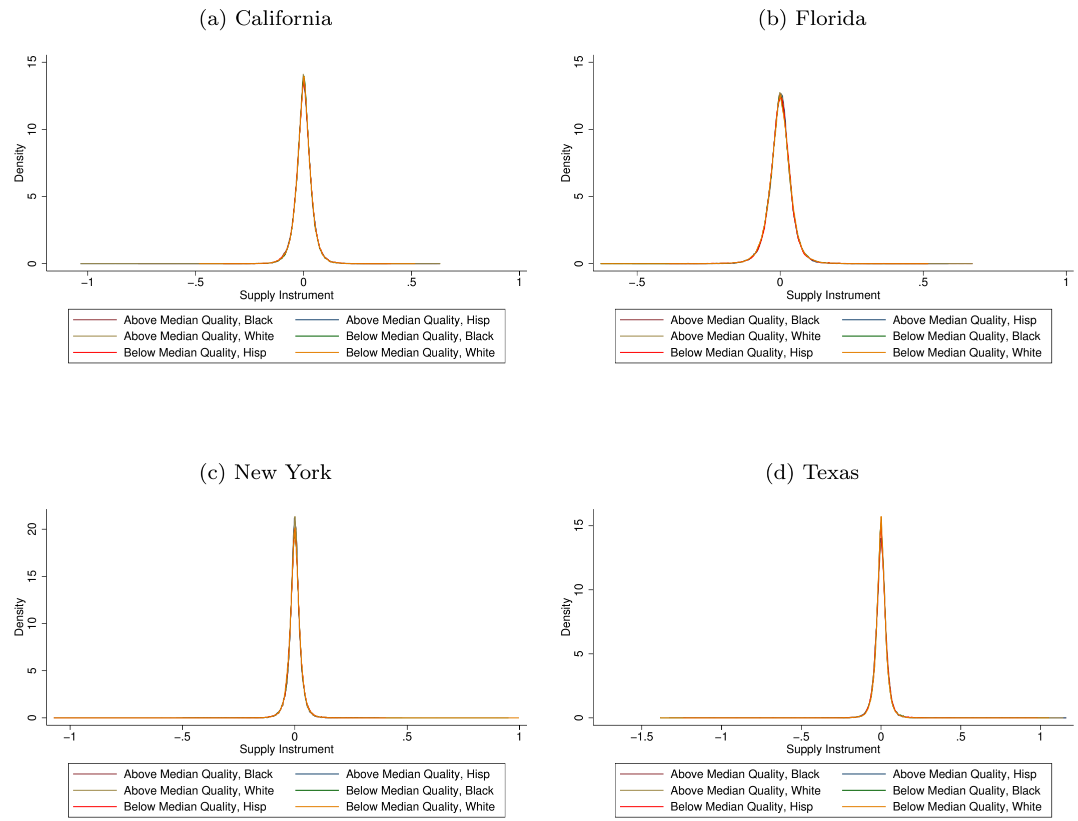


(d) Texas



Notes: This figure display kernel density plots of temporary occupancy fluctuations in the four states used for the structural estimation (defined as lagged 7-day log occupancy residualized of nursing home-month fixed effects) of nursing homes within 15 miles of each resident at their time of admission. The unit of observation is a nursing home-resident pair, and the sample is split by whether nursing homes are above-median or below-median in quality (based on LPN staffing), and the race of the resident.

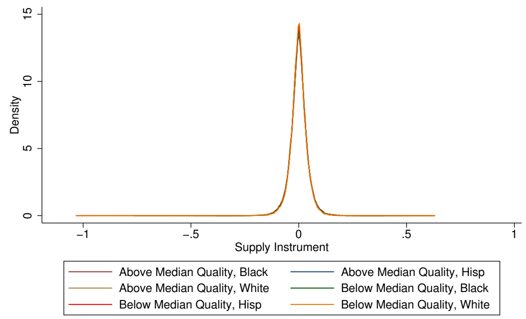
Figure A.14: Exclusion Restriction for Temporary Occupancy Fluctuations (Split by Standard Deficiencies and Race of Resident)



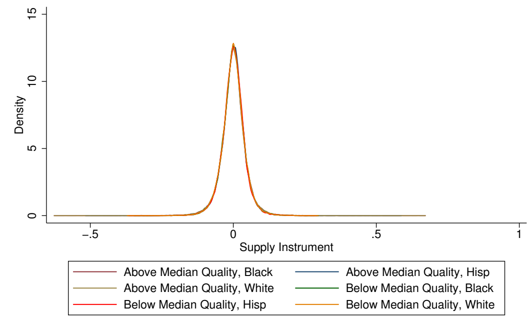
Notes: This figure display kernel density plots of temporary occupancy fluctuations in the four states used for the structural estimation (defined as lagged 7-day log occupancy residualized of nursing home-month fixed effects) of nursing homes within 15 miles of each resident at their time of admission. The unit of observation is a nursing home-resident pair, and the sample is split by whether nursing homes are above-median or below-median in quality (based on fewer standard deficiencies), and the race of the resident.

Figure A.15: Exclusion Restriction for Temporary Occupancy Fluctuations (Split by Complaint Deficiencies and Race of Resident)

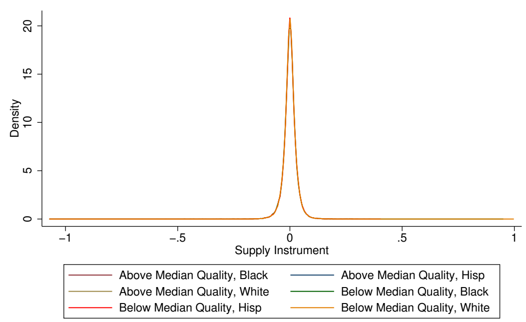
(a) California



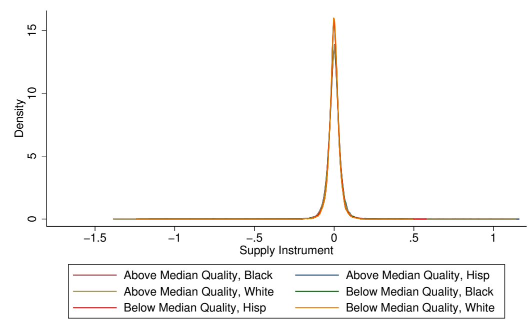
(b) Florida



(c) New York

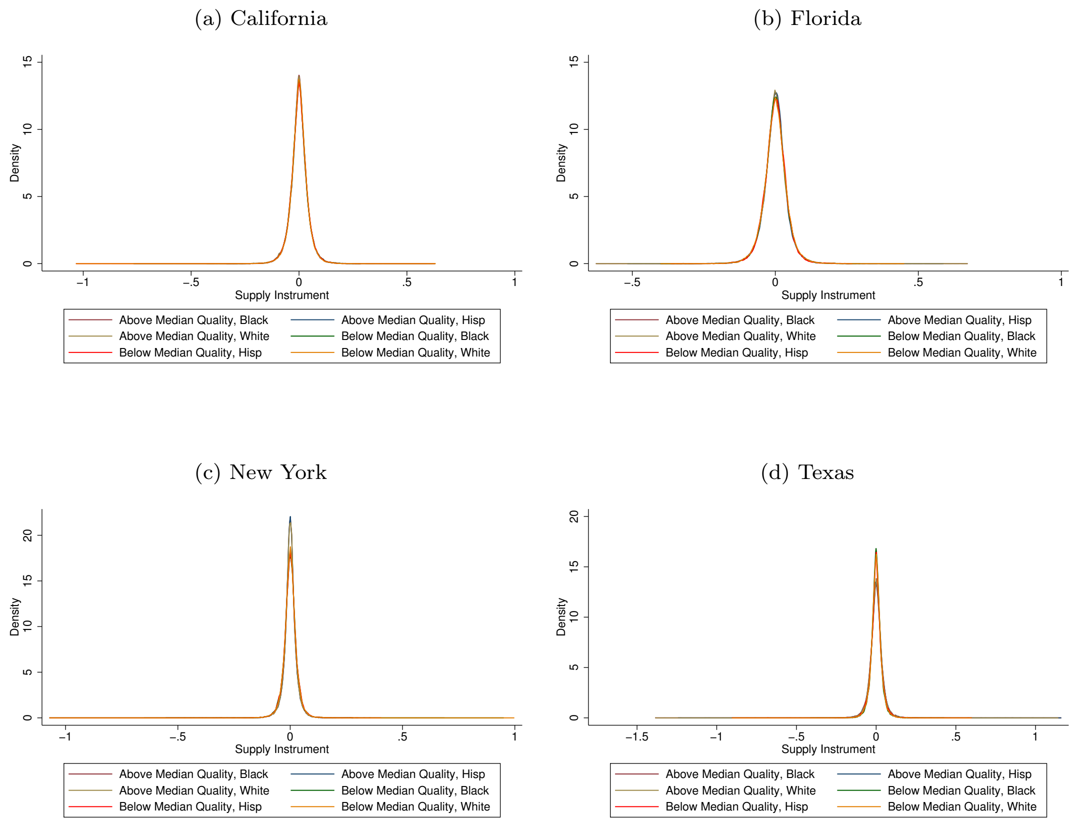


(d) Texas



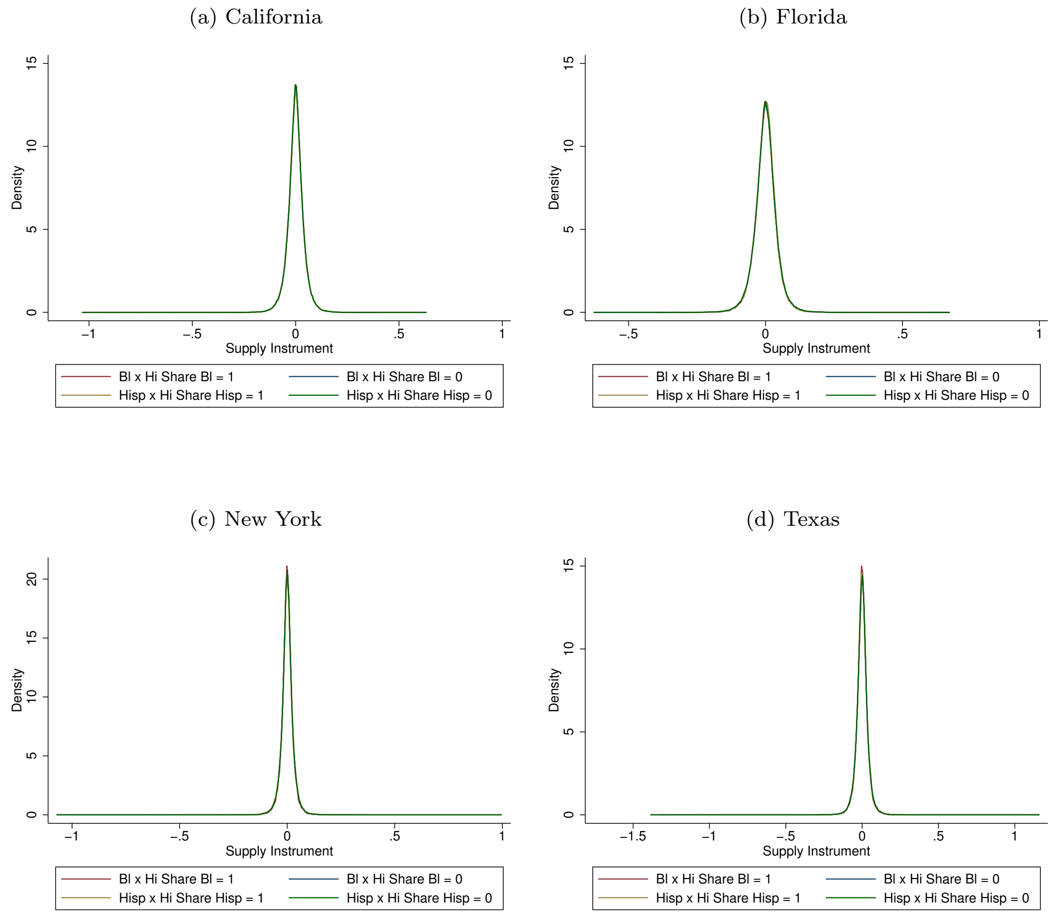
Notes: This figure display kernel density plots of temporary occupancy fluctuations in the four states used for the structural estimation (defined as lagged 7-day log occupancy residualized of nursing home-month fixed effects) of nursing homes within 15 miles of each resident at their time of admission. The unit of observation is a nursing home-resident pair, and the sample is split by whether nursing homes are above-median or below-median in quality (based on fewer complaint deficiencies), and the race of the resident.

Figure A.16: Exclusion Restriction for Temporary Occupancy Fluctuations (Split by 2009 Star Ratings and Race of Resident)



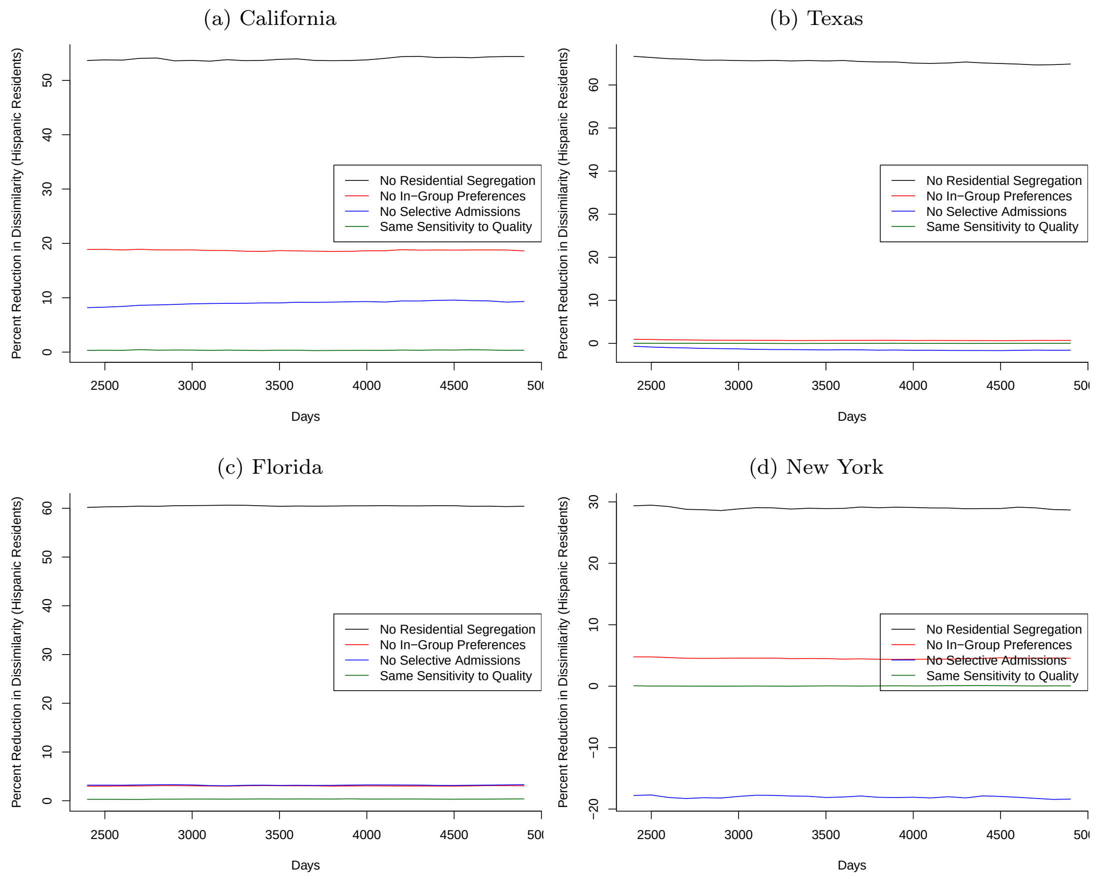
Notes: This figure display kernel density plots of temporary occupancy fluctuations in the four states used for the structural estimation (defined as lagged 7-day log occupancy residualized of nursing home-month fixed effects) of nursing homes within 15 miles of each resident at their time of admission. The unit of observation is a nursing home-resident pair, and the sample is split by whether nursing homes are above-median or below-median in quality (based on 2009 star ratings), and the race of the resident.

Figure A.17: Exclusion Restriction for Temporary Occupancy Fluctuations (Split by Share of Minority Admissions in the Past Year Interacted with Race of Resident)



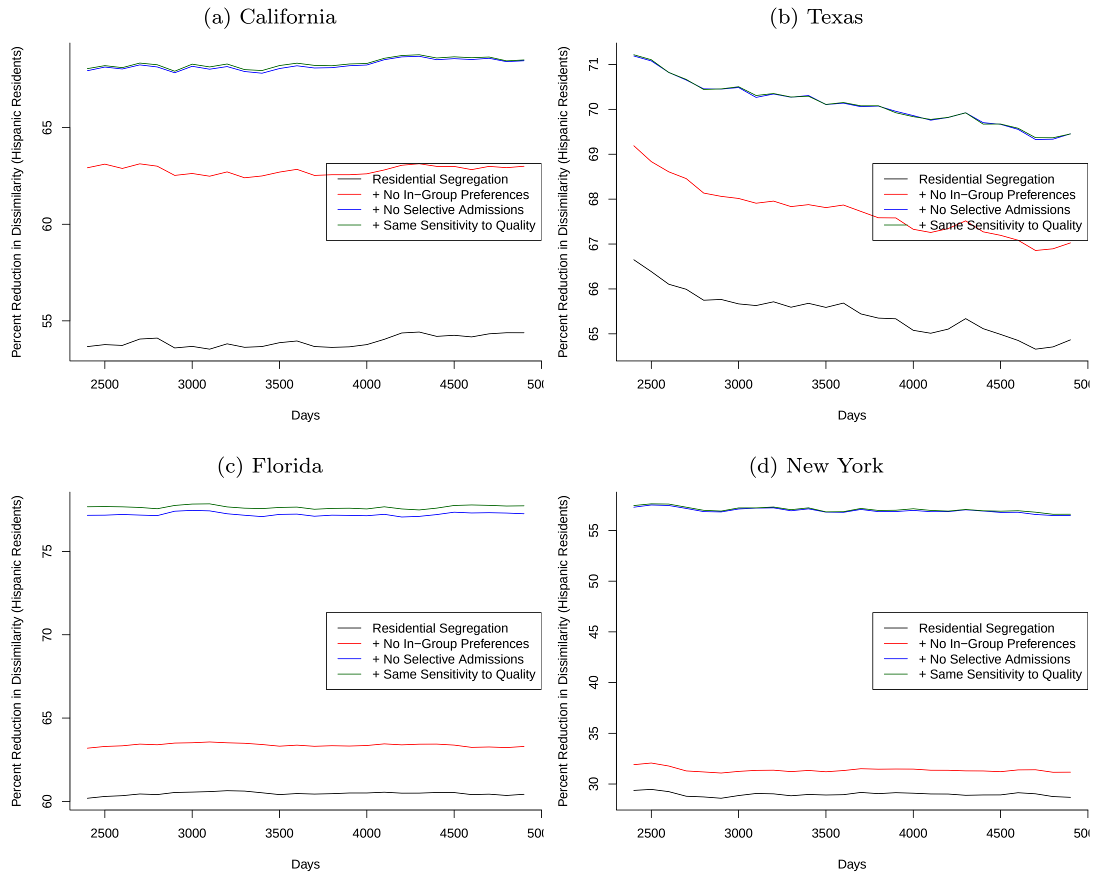
Notes: This figure display kernel density plots of temporary occupancy fluctuations in the four states used for the structural estimation (defined as lagged 7-day log occupancy residualized of nursing home-month fixed effects) of nursing homes within 15 miles of each resident at their time of admission. The unit of observation is a nursing home-resident pair, and the sample is split by whether the interaction between the share of admissions in the last 365 who are Black (or Hispanic) being above-median and whether the resident is Black (or Hispanic) is equal to one or zero.

Figure A.18: Counterfactual Segregation (Hispanic vs Non-Hispanic Residents)



Notes: These figures display the simulated average reduction in Hispanic/non-Hispanic dissimilarity index over time separately for different counterfactuals relative to the status quo simulations, over 100 replications for each counterfactual.

Figure A.19: Counterfactual Segregation with Several Explanations (Hispanic vs Non-Hispanic Residents)



Notes: These figures display the simulated average reduction in Hispanic/non-Hispanic dissimilarity index over time separately for different counterfactuals relative to the status quo simulations, over 100 replications for each counterfactual.

Table A.1: Summary Statistics for Residents by State

	California	Florida	New York	Texas
Age	77.42	77.89	77.53	77.51
	(13.10)	(12.21)	(12.88)	(12.73)
Female	0.609	0.601	0.630	0.621
	(0.488)	(0.490)	(0.483)	(0.485)
Married	0.328	0.367	0.296	0.327
	(0.469)	(0.482)	(0.457)	(0.469)
Less than High School	0.206	0.194	0.271	0.296
	(0.404)	(0.396)	(0.445)	(0.457)
High School/Some College	0.631	0.663	0.573	0.577
	(0.483)	(0.473)	(0.495)	(0.494)
At Least Bachelor's Degree	0.135	0.117	0.120	0.0953
	(0.341)	(0.322)	(0.325)	(0.294)
Medicare	0.533	0.706	0.725	0.747
	(0.499)	(0.456)	(0.447)	(0.435)
Medicaid	0.131	0.0479	0.133	0.126
	(0.337)	(0.214)	(0.340)	(0.332)
Admitted from Acute Care Hospital	0.881	0.896	0.894	0.760
	(0.324)	(0.305)	(0.308)	(0.427)
Admitted from Home	0.0869	0.0680	0.0736	0.167
	(0.282)	(0.252)	(0.261)	(0.373)
Dementia	0.231	0.220	0.234	0.271
	(0.422)	(0.414)	(0.424)	(0.444)
White	0.737	0.825	0.788	0.742
	(0.440)	(0.380)	(0.409)	(0.437)
Black	0.0685	0.0855	0.120	0.105
	(0.253)	(0.280)	(0.325)	(0.307)
Hispanic	0.109	0.0788	0.0628	0.140
	(0.311)	(0.269)	(0.243)	(0.347)
Number of Residents	773,552	816,024	677,848	500,431

Notes: This table contains summary statistics for residents who had their first stays in a nursing home between 2000 and 2010.

Table A.2: Association Between Other Measures of Nursing Home Quality and Minority Status

	RN Staffing (s.d.)		LPN Staffing (s.d.)		Fewer Standard Deficiencies (s.d.)		Fewer Complaint Deficiencies (s.d.)		2009 Star Ratings (s.d.)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Race: Black	-0.138***	-0.0714***	0.0303***	-0.0325***	-0.129***	-0.114***	-0.252***	-0.134***	-0.235***	-0.154***
	(0.0113)	(0.00492)	(0.0102)	(0.00452)	(0.0125)	(0.00442)	(0.0248)	(0.00766)	(0.0207)	(0.00849)
Race: Hispanic	-0.0726***	-0.0481***	-0.0269*	-0.0311***	-0.167***	-0.0655***	0.00766	-0.0638***	0.0448	-0.0890***
	(0.0171)	(0.00663)	(0.0161)	(0.00915)	(0.0191)	(0.00512)	(0.0186)	(0.00536)	(0.0436)	(0.0101)
Constant	0.0169**	0.00919	-0.00202	0.00456	0.0198***	0.0142***	0.0261***	0.0168***	0.0220*	0.0191*
	(0.00800)	(0.00605)	(0.00724)	(0.00548)	(0.00739)	(0.00434)	(0.00849)	(0.00605)	(0.0129)	(0.0102)
Zip Code FE		X		X		X		X		X
Num. Obs.	8,577,363	8,575,899	8,568,306	8,566,842	8,578,937	8,577,473	4,218,959	4,216,977	8,458,704	8,457,229
R-squared	0.002	0.107	0.000	0.108	0.002	0.174	0.006	0.207	0.005	0.276

Notes: The unit of observation is a resident. Standard errors are clustered at the nursing home level.

Table A.3: Characteristics of Nursing Homes in Residents' Choice Sets (By Race)

	(1)	(2)	(3)	(4)
California	All	White	Black	Hispanic
Quality Index	0.007	0.008	0.031	0.006
	(0.191)	(0.202)	(0.122)	(0.171)
Nursing Homes Within 15 Miles	48.804	39.709	78.949	68.063
	(46.789)	(39.694)	(56.849)	(57.114)
Distance to Nursing Homes	8.367	8.244	8.701	8.576
	(1.990)	(2.086)	(1.530)	(1.905)
Florida	All	White	Black	Hispanic
Quality Index	-0.014	-0.035	-0.010	0.178
	(0.271)	(0.262)	(0.266)	(0.280)
Nursing Homes Within 15 Miles	21.336	20.571	23.400	26.162
	(13.538)	(13.894)	(12.096)	(9.955)
Distance to Nursing Homes	7.975	7.931	7.748	8.578
	(1.968)	(2.014)	(1.973)	(1.329)
New York	All	White	Black	Hispanic
Quality Index	-0.056	-0.062	-0.041	-0.033
	(0.217)	(0.244)	(0.117)	(0.084)
Nursing Homes Within 15 Miles	66.770	52.924	103.172	112.745
	(53.115)	(50.030)	(41.024)	(33.720)
Distance to Nursing Homes	8.446	8.466	8.362	8.284
	(1.636)	(1.802)	(1.087)	(0.856)
Texas	All	White	Black	Hispanic
Quality Index	-0.087	-0.083	-0.140	-0.061
	(0.240)	(0.242)	(0.203)	(0.255)
Nursing Homes Within 15 Miles	15.112	14.314	20.097	14.170
	(11.696)	(11.622)	(12.239)	(10.622)
Distance to Nursing Homes	7.427	7.426	8.011	6.917
	(2.874)	(2.948)	(2.586)	(2.727)

Notes: This table contains summary statistics for nursing homes within 15 miles of each resident who had their first stays in a nursing home in the indicated state between 2008 and 2010. The unit of observation is a resident-nursing home pair.

Table A.4: Robustness Checks for Evidence on Capacity Constraints

(a) Dummy for Any New Admission as the Dependent Variable

	Any New Residents		
	(1)	(2)	(3)
Lagged 7-Day Avg. Log Occupancy	-0.206***		
	(0.00347)		
Lagged 7-Day Avg. Occupancy		-0.00588***	
		(6.71e-05)	
Lagged 7-Day Avg. Occ. Percentile			-0.00165***
			(1.47e-05)
Nursing Home-Month Fixed Effects	X	X	X
N	40,033,450	40,033,450	40,033,450
R-squared	0.205	0.169	0.169

Notes: This table shows regression results at the nursing home-day level wherein the dependent variable is a dummy for any new residents, and the independent variables are various measures of nursing home occupancy. Standard errors are clustered at the nursing home level.

(b) Flow of Residents as the Dependent Variable

	Flow of Residents		
	(1)	(2)	(3)
Lagged 7-Day Avg. Log Occupancy	-2.831***		
	(0.0374)		
Lagged 7-Day Avg. Occupancy		-0.0852***	
		(0.000430)	
Lagged 7-Day Avg. Occ. Percentile			-0.0213***
			(0.000111)
Nursing Home-Month Fixed Effects	X	X	X
N	40,033,450	40,033,450	40,033,450
R-squared	0.035	0.045	0.040

Notes: This table shows regression results at the nursing home-day level wherein the dependent variable is the flow of residents (difference between number of residents today and yesterday), and the independent variables are various measures of nursing home occupancy. Standard errors are clustered at the nursing home level.

Table A.5: Robustness Checks for Evidence on Selective Admissions

(a) Using Occupancy in Levels as the Independent Variable (Unconditional)

	Black (1)	Hispanic (2)	Post-Acute (3)
Lagged 7-Day Avg. Occupancy	-0.000421***	-0.000110***	0.00111***
	(4.90e-05)	(3.16e-05)	(6.11e-05)
Nursing Home-Month Fixed Effects	X	X	X
Number of Residents	7,106,929	7,106,929	7,106,929
R-squared	0.385	0.355	0.301

Notes: Regressions are at the resident level. Standard errors are clustered by nursing home.

(b) Using Occupancy in Levels as the Independent Variable (Conditional)

	Black (1)	Hispanic (2)	Post-Acute (3)
Lagged 7-Day Avg. Occupancy	-0.000279***	-5.08e-05	0.000741***
	(4.80e-05)	(3.15e-05)	(5.67e-05)
Nursing Home-Month Fixed Effects	X	X	X
Controls for Other Characteristics	X	X	X
Number of Residents	7,106,929	7,106,929	7,106,929
R-squared	0.385	0.355	0.301

Notes: Regressions are at the resident level, and include controls for race, Medicaid, post-acute care, dementia, age, gender, marital status, and education (as long as the variable is not the dependent variable). Standard errors are clustered by nursing home.

(c) Using Occupancy Percentile as the Independent Variable (Unconditional)

	Black (1)	Hispanic (2)	Post-Acute (3)
Lagged 7-Day Avg. Percentile	-0.000135***	-4.56e-05***	0.000365***
	(1.56e-05)	(1.02e-05)	(1.93e-05)
Nursing Home-Month Fixed Effects	X	X	X
Number of Residents	7,106,929	7,106,929	7,106,929
R-squared	0.385	0.355	0.301

Notes: Regressions are at the resident level. Standard errors are clustered by nursing home.

(d) Using Occupancy Percentile as the Independent Variable (Conditional)

	Black (1)	Hispanic (2)	Post-Acute (3)
Lagged 7-Day Avg. Percentile	-8.66e-05***	-2.47e-05**	0.000239***
	(1.54e-05)	(1.02e-05)	(1.85e-05)
Nursing Home-Month Fixed Effects	X	X	X
Controls for Other Characteristics	X	X	X
Number of Residents	7,102,426	7,102,426	7,102,426
R-squared	0.393	0.358	0.338

Notes: Regressions are at the resident level, and include controls for race, Medicaid, post-acute care, dementia, age, gender, marital status, and education (as long as the variable is not the dependent variable). Standard errors are clustered by nursing home.

Table A.6: Robustness Checks for Dynamic Panel Estimates of In-Group Preferences

Panel A: Number of Admitted Residents who are Black			
	OLS	*Anderson-Hsiao*	*Arellano-Bond*
	(1)	(2)	(3)
Number of Previous Black Admits	0.520***	2.563	0.145**
	(0.0347)	(1.803)	(0.0606)
Nursing Home Fixed Effects	X	Differenced-out	Differenced-out
County x Year Fixed Effects	X	X	X
Number of Nursing Home-Years	114,962	100,608	112,017
Panel B: Number of Admitted Residents who are Hispanic			
	OLS	*Anderson-Hsiao*	*Arellano-Bond*
	(1)	(2)	(3)
Number of Previous Hispanic Admits	0.515***	1.313***	0.532***
	(0.0193)	(0.304)	(0.0607)
Nursing Home Fixed Effects	X	Differenced-out	Differenced-out
County x Year Fixed Effects	X	X	X
Number of Nursing Home-Years	114,962	100,608	112,017

Notes: This table shows regression results at the nursing home-year level, with weights equal to the number of residents who were admitted to the nursing home during that year. The Anderson-Hsiao and Arellano-Bond specifications correspond to dynamic panel methods from Anderson and Hsiao (1982) and Arellano and Bond (1991) respectively. Standard errors are clustered at the nursing home level.

Table A.7: Effect of Short-Term Fluctuations in Racial Composition on Admissions of Different Racial Groups (Occupancy in Levels)

Panel A: White Admissions	(1)	(2)	(3)	(4)
White Occupancy	0.0675***			0.0675***
	(0.000456)			(0.000456)
Black Occupancy		0.00295***		0.00131***
		(0.000384)		(0.000391)
Hispanic Occupancy			0.00152***	0.000547
			(0.000494)	(0.000503)
Nursing Home-Month Fixed Effects	X	X	X	X
Number of Observations	40,095,480	40,095,480	40,095,480	40,095,480
R-squared	0.310	0.286	0.286	0.310
Panel B: Black Admissions	(1)	(2)	(3)	(4)
White Occupancy	0.000112**			-7.59e-05
	(4.81e-05)			(4.68e-05)
Black Occupancy		0.0643***		0.0643***
		(0.000449)		(0.000448)
Hispanic Occupancy			0.000374**	-0.000433**
			(0.000177)	(0.000170)
Nursing Home-Month Fixed Effects	X	X	X	X
Number of Observations	40,095,480	40,095,480	40,095,480	40,095,480
R-squared	0.192	0.221	0.192	0.221
Panel C: Hispanic Admissions	(1)	(2)	(3)	(4)
White Occupancy	-6.34e-05***			-0.000107***
	(2.13e-05)			(2.16e-05)
Black Occupancy		0.000180**		-0.000143*
		(7.27e-05)		(7.74e-05)
Hispanic Occupancy			0.0686***	0.0686***
			(0.000585)	(0.000585)
Nursing Home-Month Fixed Effects	X	X	X	X
Number of Observations	40,095,480	40,095,480	40,095,480	40,095,480
R-squared	0.219	0.219	0.249	0.249

Notes: This table shows regression estimates at the nursing home-day level of admissions of residents of different races on occupancy of different races, controlling for nursing home-month fixed effects. Standard errors are clustered at the nursing home level.

Table A.8: Effect of Short-Term Fluctuations in Racial Composition on Admissions of Different Racial Groups (Number of New Admissions)

Panel A: Any White Admissions	(1)	(2)	(3)	(4)
Log of White Occupancy	0.996***			1.169***
	(0.0126)			(0.0213)
Log of Black Occupancy		0.00713***		0.00235*
		(0.000779)		(0.00123)
Log of Hispanic Occupancy			0.00551***	0.00284***
			(0.000911)	(0.00109)
Nursing Home-Month Fixed Effects	X	X	X	X
Number of Observations	39,969,813	24,296,281	15,202,405	11,554,850
R-squared	0.219	0.215	0.226	0.242
Panel B: Any Black Admissions	(1)	(2)	(3)	(4)
Log of White Occupancy	0.00239***			-0.00618***
	(0.000795)			(0.00201)
Log of Black Occupancy		0.216***		0.232***
		(0.00178)		(0.00250)
Log of Hispanic Occupancy			0.000932**	0.000231
			(0.000470)	(0.000626)
Nursing Home-Month Fixed Effects	X	X	X	X
Number of Observations	39,969,813	24,296,281	15,202,405	11,554,850
R-squared	0.161	0.158	0.159	0.162
Panel C: Any Hispanic Admissions	(1)	(2)	(3)	(4)
Log of White Occupancy	0.000134			-0.00584***
	(0.000517)			(0.00172)
Log of Black Occupancy		8.46e-05		-0.00119**
		(0.000251)		(0.000523)
Log of Hispanic Occupancy			0.184***	0.191***
			(0.00173)	(0.00196)
Nursing Home-Month Fixed Effects	X	X	X	X
Number of Observations	39,969,813	24,296,281	15,202,405	11,554,850
R-squared	0.159	0.158	0.166	0.164

Notes: This table shows regression estimates at the nursing home-day level of admissions of residents of dummies for different races on log occupancy of different races, controlling for nursing home-month fixed effects. Standard errors are clustered at the nursing home level.

Table A.9: Effect of Short-Term Fluctuations in Racial Composition on Admissions of Different Racial Groups (Number of New Admissions, Occupancy in Levels)

Panel A: Any White Admissions	(1)	(2)	(3)	(4)
White Occupancy	0.0329***			0.0329***
	(0.000235)			(0.000235)
Black Occupancy		0.00140***		0.000595***
		(0.000117)		(0.000158)
Hispanic Occupancy			0.00109***	0.000617***
			(0.000190)	(0.000223)
Nursing Home-Month Fixed Effects	X	X	X	X
Number of Observations	40,095,480	40,095,480	40,095,480	40,095,480
R-squared	0.225	0.207	0.207	0.225
Panel B: Any Black Admissions	(1)	(2)	(3)	(4)
White Occupancy	-2.37e-05			-0.000162***
	(2.30e-05)			(2.71e-05)
Black Occupancy		0.0473***		0.0473***
		(0.000448)		(0.000448)
Hispanic Occupancy			0.000209*	-0.000384***
			(0.000124)	(0.000124)
Nursing Home-Month Fixed Effects	X	X	X	X
Number of Observations	40,095,480	40,095,480	40,095,480	40,095,480
R-squared	0.162	0.186	0.162	0.186
Panel C: Any Hispanic Admissions	(1)	(2)	(3)	(4)
White Occupancy	-5.26e-05***			-8.69e-05***
	(1.51e-05)			(1.64e-05)
Black Occupancy		0.000158***		-9.51e-05*
		(4.94e-05)		(5.69e-05)
Hispanic Occupancy			0.0537***	0.0537***
			(0.000857)	(0.000857)
Nursing Home-Month Fixed Effects	X	X	X	X
Number of Observations	40,095,480	40,095,480	40,095,480	40,095,480
R-squared	0.160	0.160	0.189	0.189

Notes: This table shows regression estimates at the nursing home-day level of dummies for admissions of residents of different races on occupancy of different races, controlling for nursing home-month fixed effects. Standard errors are clustered at the nursing home level.

Table A.10: Evidence on Racial Heterogeneity in Sensitivity to RN Staffing

Panel A. Dependent Variable: Indicator for Admitted Resident Being White

	(1)	(2)	(3)	(4)
RN Staffing (s.d.)	0.00219***	0.000104	0.00170	0.00131
	(0.000451)	(0.00107)	(0.00116)	(0.000866)
White Share (Previous Admits)	0.966***	0.306***	0.247***	0.176***
	(0.00179)	(0.0111)	(0.0107)	(0.00801)
Nursing Home Fixed Effects		X	X	X
County x Year Fixed Effects			X	X
Zip Code of Prior Residence				X
Number of Observations	7,689,908	7,689,788	7,689,739	7,688,164
R-squared	0.293	0.303	0.307	0.397

Panel B. Dependent Variable: Indicator for Admitted Resident Being Black

	(1)	(2)	(3)	(4)
RN Staffing (s.d.)	-0.00219***	-0.000952	-0.00179*	-0.00148**
	(0.000375)	(0.000966)	(0.00106)	(0.000744)
Black Share (Previous Admits)	0.960***	0.278***	0.245***	0.164***
	(0.00207)	(0.0133)	(0.0129)	(0.00870)
Nursing Home Fixed Effects		X	X	X
County x Year Fixed Effects			X	X
Zip Code of Prior Residence				X
Number of Observations	7,689,908	7,689,788	7,689,739	7,688,164
R-squared	0.275	0.286	0.289	0.396

Panel C. Dependent Variable: Indicator for Admitted Resident Being Hispanic

	(1)	(2)	(3)	(4)
RN Staffing (s.d.)	-0.000266	0.000864**	4.09e-05	0.000150
	(0.000238)	(0.000429)	(0.000457)	(0.000428)
Hispanic Share (Previous Admits)	0.967***	0.235***	0.175***	0.152***
	(0.00411)	(0.0170)	(0.0179)	(0.0170)
Nursing Home Fixed Effects		X	X	X
County x Year Fixed Effects			X	X
Zip Code of Prior Residence				X
Number of Observations	7,689,908	7,689,788	7,689,739	7,688,164
R-squared	0.255	0.266	0.269	0.311

Notes: This table shows regression results at the resident level. Previous share of a given race is defined as the fraction of admissions in the previous year to the nursing home that the resident is admitted to who belong to the race. Standard errors are clustered at the nursing home level.

Table A.11: Evidence on Racial Heterogeneity in Sensitivity to LPN Staffing

Panel A. Dependent Variable: Indicator for Admitted Resident Being White				
	(1)	(2)	(3)	(4)
LPN Staffing (s.d.)	-0.00190***	-0.00617***	-0.000383	-0.000781
	(0.000441)	(0.000799)	(0.000815)	(0.000692)
White Share (Previous Admits)	0.966***	0.305***	0.247***	0.176***
	(0.00172)	(0.0111)	(0.0107)	(0.00801)
Nursing Home Fixed Effects		X	X	X
County x Year Fixed Effects			X	X
Zip Code of Prior Residence				X
Number of Observations	7,689,908	7,689,788	7,689,739	7,688,164
R-squared	0.293	0.303	0.307	0.397
Panel B. Dependent Variable: Indicator for Admitted Resident Being Black				
	(1)	(2)	(3)	(4)
LPN Staffing (s.d.)	0.00201***	0.00311***	-0.000135	0.000210
	(0.000385)	(0.000670)	(0.000678)	(0.000548)
Black Share (Previous Admits)	0.961***	0.277***	0.245***	0.164***
	(0.00201)	(0.0133)	(0.0129)	(0.00871)
Nursing Home Fixed Effects		X	X	X
County x Year Fixed Effects			X	X
Zip Code of Prior Residence				X
Number of Observations	7,689,908	7,689,788	7,689,739	7,688,164
R-squared	0.275	0.286	0.289	0.396
Panel C. Dependent Variable: Indicator for Admitted Resident Being Hispanic				
	(1)	(2)	(3)	(4)
LPN Staffing (s.d.)	7.81e-06	0.00341***	0.000428	0.000535
	(0.000230)	(0.000414)	(0.000446)	(0.000418)
Hispanic Share (Previous Admits)	0.967***	0.234***	0.175***	0.152***
	(0.00409)	(0.0170)	(0.0179)	(0.0170)
Nursing Home Fixed Effects		X	X	X
County x Year Fixed Effects			X	X
Zip Code of Prior Residence				X
Number of Observations	7,689,908	7,689,788	7,689,739	7,688,164
R-squared	0.255	0.266	0.269	0.311

Notes: This table shows regression results at the resident level. Previous share of a given race is defined as the fraction of admissions in the previous year to the nursing home that the resident is admitted to who belong to the race. Standard errors are clustered at the nursing home level.

Table A.12: Evidence on Racial Heterogeneity in Sensitivity to Standard Deficiencies

Panel A. Dependent Variable: Indicator for Admitted Resident Being White

	(1)	(2)	(3)	(4)
Fewer Standard Deficiencies (s.d.)	0.00146***	0.00102***	0.000946***	0.000596***
	(0.000204)	(0.000238)	(0.000251)	(0.000217)
White Share (Previous Admits)	0.966***	0.306***	0.247***	0.176***
	(0.00174)	(0.0111)	(0.0107)	(0.00801)
Nursing Home Fixed Effects		X	X	X
County x Year Fixed Effects			X	X
Zip Code of Prior Residence				X
Number of Observations	7,689,908	7,689,788	7,689,739	7,688,164
R-squared	0.293	0.303	0.307	0.397

Panel B. Dependent Variable: Indicator for Admitted Resident Being Black

	(1)	(2)	(3)	(4)
Fewer Standard Deficiencies (s.d.)	-0.000958***	-0.000770***	-0.000751***	-0.000470***
	(0.000167)	(0.000202)	(0.000211)	(0.000174)
Black Share (Previous Admits)	0.961***	0.278***	0.245***	0.164***
	(0.00202)	(0.0133)	(0.0129)	(0.00871)
Nursing Home Fixed Effects		X	X	X
County x Year Fixed Effects			X	X
Zip Code of Prior Residence				X
Number of Observations	7,689,908	7,689,788	7,689,739	7,688,164
R-squared	0.275	0.286	0.289	0.396

Panel C. Dependent Variable: Indicator for Admitted Resident Being Hispanic

	(1)	(2)	(3)	(4)
Fewer Standard Deficiencies (s.d.)	-0.000549***	-0.000273**	-0.000208	-0.000133
	(0.000125)	(0.000131)	(0.000145)	(0.000139)
Hispanic Share (Previous Admits)	0.967***	0.235***	0.175***	0.152***
	(0.00411)	(0.0170)	(0.0179)	(0.0170)
Nursing Home Fixed Effects		X	X	X
County x Year Fixed Effects			X	X
Zip Code of Prior Residence				X
Number of Observations	7,689,908	7,689,788	7,689,739	7,688,164
R-squared	0.255	0.266	0.269	0.311

Notes: This table shows regression results at the resident level. Previous share of a given race is defined as the fraction of admissions in the previous year to the nursing home that the resident is admitted to who belong to the race. Standard errors are clustered at the nursing home level.

Table A.13: Evidence on Racial Heterogeneity in Sensitivity to Complaint Deficiencies

Panel A. Dependent Variable: Indicator for Admitted Resident Being White				
	(1)	(2)	(3)	(4)
Fewer Complaint Deficiencies (s.d.	0.00128***	0.00571***	0.00245***	0.00195***
	(0.000268)	(0.000370)	(0.000429)	(0.000336)
White Share (Previous Admits)	0.966***	0.298***	0.246***	0.175***
	(0.00172)	(0.0109)	(0.0106)	(0.00799)
Nursing Home Fixed Effects		X	X	X
County x Year Fixed Effects			X	X
Zip Code of Prior Residence				X
Number of Observations	7,689,908	7,689,788	7,689,739	7,688,164
R-squared	0.293	0.303	0.307	0.397
Panel B. Dependent Variable: Indicator for Admitted Resident Being Black				
	(1)	(2)	(3)	(4)
Fewer Complaint Deficiencies (s.d.	-0.00146***	-0.00424***	-0.00206***	-0.00166***
	(0.000242)	(0.000351)	(0.000402)	(0.000306)
Black Share (Previous Admits)	0.961***	0.272***	0.244***	0.164***
	(0.00201)	(0.0130)	(0.0128)	(0.00868)
Nursing Home Fixed Effects		X	X	X
County x Year Fixed Effects			X	X
Zip Code of Prior Residence				X
Number of Observations	7,689,908	7,689,788	7,689,739	7,688,164
R-squared	0.275	0.286	0.289	0.396
Panel C. Dependent Variable: Indicator for Admitted Resident Being Hispanic				
	(1)	(2)	(3)	(4)
Fewer Complaint Deficiencies (s.d.	5.13e-05	-0.00178***	-0.000433**	-0.000334**
	(0.000133)	(0.000145)	(0.000168)	(0.000161)
Hispanic Share (Previous Admits)	0.967***	0.232***	0.175***	0.152***
	(0.00409)	(0.0171)	(0.0179)	(0.0170)
Nursing Home Fixed Effects		X	X	X
County x Year Fixed Effects			X	X
Zip Code of Prior Residence				X
Number of Observations	7,689,908	7,689,788	7,689,739	7,688,164
R-squared	0.255	0.266	0.269	0.311

Notes: This table shows regression results at the resident level. Previous share of a given race is defined as the fraction of admissions in the previous year to the nursing home that the resident is admitted to who belong to the race. Standard errors are clustered at the nursing home level.

Table A.14: Estimates of Residents' Preferences and Selective Admissions by Nursing Homes (Fixed Effects in Utility)

	California	Florida	New York	Texas
Residents' Preferences	(1)	(2)	(3)	(4)
Utility (Distance to Facility)	-0.173***	-0.19***	-0.209***	-0.176***
	(0.001)	(0.001)	(0.001)	(0.001)
Utility (Previous Share Black)	-0.42***	-0.398***	-0.642***	-0.55***
	(0.062)	(0.082)	(0.072)	(0.092)
Utility (Black x Previous Share Black)	1.711***	2.342***	2.119***	2.375***
	(0.035)	(0.051)	(0.025)	(0.046)
Utility (Previous Share Hispanic)	-0.609***	-0.826***	-0.635***	-0.896***
	(0.061)	(0.104)	(0.088)	(0.091)
Utility (Hispanic x Previous Share Hispanic)	2.331***	2.102***	2.303***	2.245***
	(0.033)	(0.069)	(0.06)	(0.06)
Utility (Quality Index)	0.018***	-0.009*	0.022***	0.062***
	(0.004)	(0.006)	(0.006)	(0.008)
Utility (Quality Index x Black)	-0.021*	-0.148***	-0.063***	-0.082***
	(0.011)	(0.013)	(0.009)	(0.015)
Utility (Quality Index x Hispanic)	-0.041***	-0.132***	-0.057***	-0.135***
	(0.008)	(0.016)	(0.01)	(0.016)
Nursing Homes' Admission Policies				
Occupancy	-3.645***	-4.463***	-4.722***	-3.219***
	(0.223)	(0.375)	(0.322)	(0.29)
Race (Black)	-0.934***	-0.864***	-0.504***	-0.827***
	(0.025)	(0.04)	(0.042)	(0.039)
Race (Hispanic)	-0.46***	-0.952***	-0.352***	-0.123**
	(0.018)	(0.035)	(0.049)	(0.048)
Nursing Home Fixed Effects in Utility	X	X	X	X

Notes: This table shows estimates of the structural model using Gibbs sampling. A burn-in period corresponding to the first half of the chain was used. The supply-side equation includes an intercept, the quality index, and a dummy for the year 2010.

Table A.15: Estimates of Residents' Preferences and Selective Admissions by Nursing Homes (Multiple Quality Measures)

	California	Florida	New York	Texas
Residents' Preferences	(1)	(2)	(3)	(4)
Utility (Distance to Facility)	-0.165***	-0.177***	-0.19***	-0.164***
	(0.001)	(0.001)	(0.001)	(0.001)
Utility (Previous Share Black)	-1.029***	-1.546***	-1.185***	-1.464***
	(0.021)	(0.021)	(0.014)	(0.025)
Utility (Black x Previous Share Black)	1.836***	1.956***	1.714***	2.065***
	(0.039)	(0.037)	(0.021)	(0.044)
Utility (Previous Share Hispanic)	-1.073***	-1.086***	-1.141***	-1.365***
	(0.02)	(0.028)	(0.022)	(0.033)
Utility (Hispanic x Previous Share Hispanic)	2.178***	1.241***	2.008***	1.471***
	(0.034)	(0.034)	(0.03)	(0.058)
Utility (RN Staffing)	0.144***	0.186***	0.117***	0.414***
	(0.006)	(0.004)	(0.003)	(0.016)
Utility (RN Staffing x Black)	-0.346***	-0.086***	-0.086***	-0.083**
	(0.019)	(0.014)	(0.007)	(0.035)
Utility (RN Staffing x Hispanic)	-0.048***	-0.055***	-0.08***	-0.191***
	(0.015)	(0.012)	(0.011)	(0.037)
Utility (LPN Staffing)	0.071***	0.193***	-0.076***	0.173***
	(0.003)	(0.006)	(0.013)	(0.012)
Utility (LPN Staffing x Black)	-0.125***	-0.191***	-0.332***	-0.1***
	(0.011)	(0.017)	(0.016)	(0.031)
Utility (LPN Staffing x Hispanic)	-0.044***	-0.171***	-0.329***	-0.303***
	(0.008)	(0.015)	(0.017)	(0.029)
Utility (Fewer Standard Deficiencies)	-0.03***	-0.015***	0.044***	0.006*
	(0.002)	(0.003)	(0.003)	(0.004)
Utility (Fewer Standard Deficiencies x Black)	0.027***	-0.065***	-0.032***	0.03***
	(0.008)	(0.008)	(0.006)	(0.009)
Utility (Fewer Standard Deficiencies x Hispanic)	-0.006*	-0.073***	-0.037***	0.033***
	(0.005)	(0.008)	(0.007)	(0.009)
Utility (Fewer Complaint Deficiencies)	-0.006**	-0.075***	-0.023***	-0.017***
	(0.003)	(0.003)	(0.005)	(0.004)
Utility (Fewer Complaint Deficiencies x Black)	-0.038***	-0.12***	-0.048***	-0.13***
	(0.01)	(0.01)	(0.007)	(0.009)
Utility (Fewer Complaint Deficiencies x Hispanic)	-0.024***	-0.116***	-0.03***	-0.111***
	(0.007)	(0.009)	(0.008)	(0.01)
Nursing Homes' Admission Policies				
Occupancy	-3.31***	-3.981***	-4.342***	-3.576***
	(0.307)	(0.301)	(0.166)	(0.226)
Race (Black)	-1.111***	-0.799***	-0.725***	-0.7***
	(0.017)	(0.046)	(0.042)	(0.032)
Race (Hispanic)	-0.122***	-0.782***	-0.718***	-0.481***
	(0.04)	(0.024)	(0.059)	(0.025)

Notes: This table shows estimates of the structural model using Gibbs sampling. A burn-in period corresponding to the first half of the chain was used. The supply-side equation includes an intercept, RN staffing, LPN staffing, standard deficiencies, complaint deficiencies, and a dummy for the year 2010.