NBER WORKING PAPER SERIES

RATIONING BY RACE

Manasvini Singh
Atheendar Venkataramani

Working Paper 30380
http://www.nber.org/papers/w30380

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2022, Revised April 2026

We thank the editor and four anonymous referees for their extremely thoughtful comments. We gratefully acknowledge feedback from seminar participants at the NBER Summer Institute, NBER Conference on Racial and Ethnic Health Disparities, the Health Economics Workshop at NYU Stern School of Business, AEA Annual Meeting, Becker Friedman Institute at the University of Chicago, Harvard University (Kennedy), University of Pennsylvania, Stanford University, Duke University, Carnegie Mellon University, University of Texas at Austin, and University of Virginia. For helpful comments and suggestions, we thank David Asch, Bocar Ba, Joseph Benitez, Bernard Black, Jacob Bor, Kate Bundorf, Kitt Carpenter, Ben Castleman, David Chan, Amitabh Chandra, Paula Chatterjee, David Cutler, Janet Currie, Michael Darden, Kit Delgado, Matthew Eisenberg, Jeremy Goldhaber-Fiebert, Jose Fernandes, Ryan Greyson, Ezra Golberstein, Jevay Grooms, Scott Halpern, Noah Hammarlund, Kandice Kapinos, Mireille Jacobson, Hedwig Lee, Jeff Linder, Elizabeth Linos, Judith Long, Dan Ly, Amol Navathe, Ziad Obermeyer, Maya Rossin-Slater, Rosalie Pacula, Carol Propper, Seth Richards-Shubik, Christopher Ruhm, Hannes Schwandt, Peter Schwardmann, Hummy Song, Zirui Song, Sebastian Tello-Trillo, Rachel Werner, Morgan Williams, Jr., Michelle White, Derek Wu, and Jacob Zureich. We thank Ashvin Gandhi for the title of this paper. This paper substantially updates a previous version, titled “Capacity strain and racial disparities in mortality.” The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Rationing by Race
Manasvini Singh and Atheendar Venkataramani
NBER Working Paper No. 30380
August 2022, Revised April 2026
JEL No. I10, I14, J15

ABSTRACT

We document how deepening resource scarcity results in rationing on the basis of race in a high-stakes setting: health care. Using detailed, time-stamped data on 107,000 inpatient admissions to a large health system, we find that in-hospital mortality increases for Black, but not White, patients as hospitals reach capacity. These findings are not explained by differential patient selection. We identify rationing by wait times as a mechanism, documenting that sicker Black patients wait longer for care than healthier White patients at almost all capacity levels. Text analysis of unstructured clinical notes reveals rationing of provider effort as another mechanism.

Manasvini Singh
Carnegie Mellon University
msingh01@cmu.edu

Atheendar Venkataramani
University of Pennsylvania
Perelman School of Medicine
Department of Medical Ethics and
Health Policy
and NBER
atheenv@pennmedicine.upenn.edu

1 Introduction

Decision-makers seek to ration scarce resources by need, willingness to pay (WTP), or other dimensions consistent with social values (Akbarpour et al., 2024; Sen, 1982; Weitzman, 1977). However, when resource scarcity deepens, they may instead ration resources in socially suboptimal ways, such as on the basis of group identity. Discrimination in resource allocation is well documented in the United States (Arrow, 1998; Darity, 2022; Darity & Mason, 1998), and lab-based research suggests that resource scarcity and competition over finite resources increases discriminatory behaviors and beliefs (Antunes et al., 2023; Berkebile-Weinberg, Krosch, & Amodio, 2022; Haushofer et al., 2023; Krosch, Tyler, & Amodio, 2017). We investigate whether resource scarcity increases rationing on the basis of *race* in a highly consequential, real-world setting: health care.

Racial discrimination in contemporary health care is widespread (Institute of Medicine, 2003), reflecting long-standing assumptions about genetic differences and cultural practices (Eli, Logan, & Miloucheva, 2023) and a history of overt medical mistreatment of racial minority populations (Alsan & Wanamaker, 2018). Contemporary research documents multiple mechanisms underpinning discrimination, including taste-based and statistical bias among providers (Schulman et al., 1999; Stepanikova, 2012), algorithmic bias in clinical decision tools (Obermeyer et al., 2019), social distance between providers and patients (Alsan, Garrick, & Graziani, 2019; Schwab & Singh, 2023), and structural disadvantages related to geography and insurance coverage (Chandra & Skinner, 2003; Doyle, 2005). We hypothesize that as resource, personnel, and time constraints on health care systems and providers become binding, such mechanisms may generate health care rationing on the basis of race rather than medical need, harming patient outcomes.

To explore this hypothesis, we use detailed, time-stamped electronic health records on over 107,000 patient admissions spanning 2015–2018 across two hospitals within a large urban academic hospital system in the southeast United States. We leverage plausibly exogenous variation in hourly hospital capacity as our measure of resource scarcity. As hospitals approach maximum capacity, increasing patient demands can strain critical resources, such as beds, personnel, diagnostic tools, treatments, and provider bandwidth (Hoe, 2022; Song et al., 2020). Such so-called capacity strain (Arogyaswamy et al., 2021) may exacerbate taste-based discrimination among providers or reduce the time and cognitive bandwidth available to them to accurately evaluate patient acuity and deliver suitable care (Johnson et al., 2016; Stepanikova, 2012). We construct hospital-specific deciles of strain, as measured by a census of patients in the hospital in a given hour, to capture potential nonlinearities in the effects of strain and facilitate comparisons at similar strain levels across hospitals that differ across measures of operational capacity, such as staffing, clinical resources, and local demand mix.

We find that in-hospital mortality among Black patients increases as hospitals become more strained – a pattern not documented for White patients. With an average mortality rate of 1.6% in our sample, Black patients arriving to a hospital at its highest decile of strain (in our sample, a bed occupancy rate of 91–95%) are 0.7 percentage points (pp) likelier to die than White patients, while being 0.2 pp likelier to die at lower deciles of strain. These findings are robust to our defining hospital strain in an alternative ways and employing both sparser and richer sets of covariates. The strain-related increases in racial mortality gaps are driven by patients with complex, difficult-to-assess medical needs. In contrast, patients admitted with more objective symptoms, for which initial care tasks are highly protocolized, do not exhibit significant racial disparities in mortality. These findings suggest that when time, bandwidth, and resources are constrained, providers and routine care processes fall short in diagnosing and treating patients, particularly under clinical uncertainty, with disproportionate consequences for Black patients.

We next assess several alternative explanations that could account for our findings, such as shifts in patient composition. These analyses also assess the validity of our identifying assumption of conditional independence: that any observed and unobserved differences between Black and White patients in our sample that may affect mortality outcomes remain constant across different levels of capacity strain. This assumption is supported by the literature demonstrating that the day-to-day variation in hospital strain – let alone the hour-to-hour variation that we can uniquely leverage given our access to proprietary time-stamped data (Neprash et al., 2021; Song et al., 2020)– is difficult to predict. Even when hospital strain can be anticipated, its adverse effects remain challenging to mitigate given the limited flexibility and substitutability of many hospital resources, an insight that supports a common identification strategy in the literature on hospital capacity, which is to assume that, conditional on the inclusion of time-related fixed effects in the estimation, capacity strain is an exogenous regressor (Hoe, 2022; Song et al., 2020).

We first rule out differential patient selection either *into* or *out of* our inpatient sample as strain increases. Our estimates would be biased if, for example, admitted Black but not White patients became sicker as strain rose. In contrast, under increasing strain, there are no differential changes by race in admitted patients' key covariates, probability of admission from the emergency department, mode of arrival to the hospital, or clinical themes that we identify in unstructured documentation by means of machine-learning (ML) methods. Nor do we find differential changes by race in the probability of *discharge* (e.g., to hospice or another hospital) as strain increases.

We next present direct evidence that race-based rationing of hospital resources is a likely mechanism behind the racial mortality gap. While there are multiple margins along which hospital resources may be rationed, we focus on two that are measurable in our data and conceptually consequential for in-hospital mortality: (i) wait time for an inpatient bed and (ii) provider effort as reflected by patterns in the unstructured clinical documentation. Wait time is a well-documented margin of rationing in health care and other service settings (Chen et al., 2022; Holt & Vinopal, 2023; Lu, Hanchate, & Paasche-Orlow, 2021; Martin & Smith, 1999). In our setting, we observe a pattern similar to that of our mortality findings: Black patients experience greater increases in wait times for beds than do White patients as capacity strain increases. In addition, providers appear to increasingly rely on patient race rather than medical need to assign inpatient beds as constraints bind: At high strain, wait times for high- versus low-need patients become indistinguishable, and significant disparities between Black and White patients persist. Indeed, at all levels of capacity strain, *high-need* Black patients wait longer for an inpatient hospital bed than *low-need* White patients.

Next, we find evidence suggesting race-based rationing of provider effort – a scarce resource viewed as a key driver of racial disparities in health care (King et al., 2023). As effort is not directly measured in electronic medical records, we follow other studies (Chan, 2018) in inferring it from patterns it leaves in data. We examine unstructured, free-text entries in the electronic health record of the reason for admission, a field filled out by a triage provider around the time of hospital arrival that plays an important role in setting the entire course of care (Schrader & Lewis, 2013). We apply several ML and text analysis methods to the text data from this field to create an index of provider effort from five features: character count, average word length, use of adjectives, subjectivity, and polarity – all dimensions that reflect effort (Goddu et al., 2018; Kennedy & Levin, 2008; Tausczik & Pennebaker, 2010; Yadav, Prabhu, & Chandy, 2007). In general, medical documentation consistently reflects lower provider effort for Black than for White patients. Mirroring the patterns in mortality and wait times, effort disproportionately declines for Black patients when hospitals reach the highest decile of strain. Finally, we discuss how our findings align with economic models of discrimination, such as taste-based, statistical, and systemic discrimination that is amplified in health care settings.

We contribute significantly to two foundational areas of economic inquiry—*scarcity* and *discrimination*—by revealing novel and consequential linkages between the two.

First, our findings show that when resource scarcity deepens, underlying discrimination comes to the fore in the form of racially biased allocation practices. This insight may explain why discrimination is a defining feature of some settings but not others. Our findings reconcile a range of labor market patterns: for example, that migrant workers were three times likelier than nonmigrant workers to be laid off in the industries most disrupted by the COVID-19 pandemic (Auer, 2022) and that applicants with foreign-sounding names are as likely as those with local-sounding names to be called back for job interviews in tight labor markets but face a significant disadvantage in loose labor markets (Baert et al., 2015). More generally, this paper also helps us understand disparities in the impacts of policies or shocks that increase scarcity. For example, research shows racial disparities in a range of economic and health outcomes manifest in the face of large-scale stressors such as economic shocks, natural disasters, pandemics, and labor shortages (Anderson, Crost, & Rees, 2020; Beck & Tolnay, 1990; Klein et al., 2023; Stoye & Warner, 2023). Similarly, policy-driven cuts to funding in sectors such as education disproportionately affect racial and ethnic minority students (Jackson, Wigger, & Xiong, 2021). Correspondingly, we make a critical methodological point: Resource shocks can serve as natural experiments in which to test for discrimination, particularly in real-world settings where direct measures are difficult to obtain.

Further, this relationship between scarcity and discrimination suggests that policies aimed at alleviating resource scarcity can be viable as interventions for reducing discriminatory practices. For instance, consider current and historical discrimination against Black home buyers – a key factor in persistent racial wealth gaps (Bayer, Ferreira, & Ross, 2018; Charles & Hurst, 2002; Rugh & Massey, 2010). Our study suggests that increasing housing supply would not only accomplish the primary goals favored by proponents of this policy (e.g., alleviating homelessness, helping younger individuals build assets (Glaeser & Gyourko, 2018; Green & White, 1997)) but could also reduce discriminatory rationing of resources in housing markets, mitigating racial disparities in homeownership and wealth accumulation. This reasoning could even apply to a range of other settings: Increasing available resources for schools, for example, may mitigate racial disparities in student outcomes (Card & Krueger, 1996), interventions to increase cognitive bandwidth among police officers may mitigate racial disparities in arrests and the use of force (Dube, MacArthur, & Shah, 2023), and interventions to improve available capacity in hospitals may help reduce disparities in health care outcomes (Vohra et al., 2023; White, Villarroel, & Hick, 2021).

Second, the literature on rationing under scarcity has largely focused on how agents ration goods using price (e.g., willingness-to-pay) or non-price (e.g., lotteries or queues) mechanisms (Breza, Kaur, & Shamdasani, 2021; Glaeser & Luttmer, 2003; Stiglitz & Weiss, 1981). The criteria underlying these mechanisms are, at least formally, *economic* – factors such as creditworthiness, wage rigidity, etc. Our findings support that a *social* factor – group identity – can operate as a rationing criterion in its own right, and should be integrated into positive theories of rationing to explain skewed outcomes in sectors where the stakes for equitable distribution are high.

Third, the literature on the distribution of resources under scarcity has been primarily concerned with the outcomes of rationing at the macro- level, i.e., how governments or markets allocate scarce resources through mechanisms that may generate inequality (Stiglitz, 2012) and conflict (André & Platteau, 1998). In contrast, this study delves into the micro-level dynamics of scarcity: we highlight how *individual* discriminatory preferences shape the allocation of scarce resources at the point of service delivery, with consequences that deepen as capacity tightens.

The results of this paper also speak specifically to the health care domain, by contributing to the literature on discrimination by illustrating its consequences for health. Studies focus on identifying discrimination in specific diagnostic or treatment decisions (Schulman et al., 1999; Stepanikova, 2012) or in process measures such as medical documentation or wait times (Goddu et al., 2018; Lu, Hanchate, & Paasche-Orlow, 2021). Our study moves further by documenting the impacts of such discrimination in a real-world setting on an important dimension of health – mortality risk – and identifying the conditions under which these impacts materialize. We note that our findings are derived from a sample of admissions that occurred prior to the COVID-19 pandemic. Race-based rationing may in fact have been far more consequential during the COVID-19 pandemic – a time of record capacity strain especially in hospitals serving Black patients (Vohra et al., 2023), extreme provider burnout (Kadri et al., 2021), and staggering racial disparities in health access and outcomes (Price-Haywood et al., 2020). The COVID-19 pandemic also gave rise to vociferous discussions over whether race-based rationing, as a means of targeting resources to patients in need, was ethical or legal (Jost, 2022). Our results show that, whatever the views on de jure rationing, de facto rationing by race appears typical of care settings under routine hospital stressors – with stark consequences.

2 Setting, Data, and Measures

2.1 Setting and data

For our main analysis, we draw on comprehensive electronic medical record (EMR) data for 107,224 inpatient admissions at two hospitals in a large, well-regarded academic health system; the data for Hospital 1 cover June 2015–June 2017 and those from Hospital 2 June 2015–June 2018. These are teaching hospitals in a southeastern US city known for its sizable Black population. Both offer medical, surgical, intensive care, and level-I trauma services, and one is highly ranked nationally in multiple specialty services. They are both typical of settings where Black patients receive health care, as large teaching hospitals are disproportionately represented among the hospitals that care for Black patients (Burke et al., 2017). The data include timestamps up to the second, documenting an inpatient's entire journey through each hospital wing from arrival to discharge. They also include mode of arrival to the hospital (e.g., ambulance, walk-in), vital signs on admission, patient sociodemographics and insurance status, physician-of-record identifiers, a comprehensive list of patient diagnoses and procedures, discharge disposition, and text from a provider-input field documenting the reason for admission.

Our sample includes patients admitted to the hospital through the emergency department (ED, $\sim 47\%$ of admissions) and those directly admitted from outpatient and specialty clinics or transfers from other facilities. We consider these groups together because patients admitted through both ED and non-ED pathways can be high acuity on arrival or have the potential to become very ill during the course of the hospitalization (Kocher, Dimick, & Nallamothu, 2013). In addition, we leverage a separate dataset comprising all ED visits to the hospitals during the study period (the "ED Sample"). This dataset includes patients discharged home without being admitted, who do not appear in our main sample. While these data are less detailed than our inpatient data, they do include triage scores for all patients, which reflect provider assessments of medical need at ED arrival. We use these data to assess differential selection into the inpatient sample with increasing hospital strain.

2.2 Construction of key measures

Patient race: We focus on non-Hispanic Black (hereafter, "Black") and non-Hispanic White ("White") patients. Our sample includes all admissions of patients in these racial categories; we lack sufficient power to analyze outcomes for patients of other races or ethnicities. As is typical in medical settings, race is either self-reported or recorded by staff. Our data, similarly to most EMR data, do not distinguish the type of reporting. Regardless, measurement error in patient racial classification is unlikely to affect our analysis: Errors in race attribution are possible but infrequent (Cook, McGuire, & Zaslavsky, 2012).

Black patients account for 60% of the admissions to the two study hospitals, consistent with the city's population being predominantly Black. Table 1 highlights differences between the Black and White patients in our sample. Compared to the White patients, the Black patients are on average younger (52 vs. 59 years) and likelier to be female (65% vs. 50%). Despite being younger, the Black patients have similar comorbidity burdens and similar in-hospital death rates (both corresponding to approximately 2% of patient admissions). They also wait over 2 hours longer on average than White patients for an inpatient hospital bed.

Hospital strain: We proxy resource scarcity with a measure of hospital strain. Our strain measure is based on the total number of inpatients in the hospital at any given hour. We choose this census-based measure given evidence from the clinical literature of its strength in predicting health outcomes (Kohn et al., 2019). We assign hospital strain to each patient at time of hospital *arrival*, not *admission*, as the time of ward admission is endogenous but the arrival time is plausibly less so. The time around arrival is likely among the most critical moments during a hospital stay, as key clinical tasks such as initial history-taking, evaluation, diagnostic testing, and bed assignment occur early in the hospitalization. Strain-induced disruptions during these initial processes can cascade, causing cumulative deficits in resource allocation. Because patient health is inherently path dependent, early misallocation of resources due to high strain may result in adverse outcomes that are difficult or impossible to reverse later in the stay.

We measure hospital strain using hourly inpatient occupancy. For each hospital separately, we construct the distribution of hourly occupancy over the sample period and assign each hour a within-hospital percentile rank, where the 1st and 100th percentiles correspond to the lowest and highest occupancy levels, respectively. For our main figures, we aggregate these percentiles into occupancy deciles, with decile 1 corresponding to percentiles $1-10$ and decile 10 corresponding to percentiles $91-100$. Each patient is then assigned the occupancy decile corresponding to the hour closest to their arrival time. For example, a patient arriving when occupancy is in the 95th percentile of that hospital's historical distribution would be assigned to occupancy decile 10.

We use within-hospital percentiles rather than absolute occupancy (e.g., percent of beds filled) to capture how strained a hospital is relative to its typical operating conditions. The level of occupancy at which strain emerges depends on hospital-specific factors – such as staffing flexibility, patient flow, and resources – so identical absolute occupancy rates can correspond to very different levels of operational stress across hospitals.[1] By normalizing census within each hospital, our measure makes high-decile observations comparable across hospitals, even when their absolute capacity levels differ.

In our main models, we operationalize strain both as an indicator denoting the 10th decile as high strain and deciles 1–9 pooled together as low strain and as a series of indicators denoting each decile. In robustness checks, we consider alternative percentile cutoffs for defining high and low strain. The mean proportion of inpatient beds filled in the first decile of strain across both hospitals ranges from 69% to 78%; at the tenth decile, the mean range of filled beds is 91% to 95% (Table A.1). Moreover, the distributions for Black and White patients are nearly identical in terms of hour of hospital arrival and average hospital strain at arrival time (Figure I), suggesting that Black and White patients do not arrive at uniquely different times of hospital strain.

<u>Patient medical need</u>: We use the Elixhauser mortality index as our measure of patient medical need (acuity). This index predicts a patient's mortality risk from a weighted combination of 29 chronic

[1] For example, a critical access hospital with less than 25 beds and a small, fixed number of staff may reach capacity strain when 30% of beds are filled, while a large teaching hospital used to large patient flows and with access to flexible staffing may not experience strain until over 90% of beds are full. Consistent with this contention, in Figure A.2, we demonstrate how average mortality starts rising at similar hospital-specific percentiles of capacity strain in our two study hospitals (near 80 percentile), even though the specific percentiles imply different percentages of beds filled (Table A.1).

conditions and has been heavily validated and widely used in medical research (Elixhauser et al., 1998; Moore et al., 2017). "High-need" and "low-need" patients are those above and below the sample median value for the Elixhauser mortality index, respectively. Figure II provides evidence that the mortality index scores are an appropriate proxy for medical need for both Black and White patients. Panel (a) plots the distribution of mortality scores and relationship with the probability of in-hospital death by patient race. The proportion of patients dying in the hospital rises with the scores and does so similarly for both Black and White patients.

To validate that the scores capture dimensions of medical need at least partially recognized by clinicians, Figure II Panel (b) plots for the sample of inpatients admitted through the ED each admitted patient's score against the ED triage score. Triage scores (reverse coded such that higher values indicate higher need) range from 1 to 5. These scores are assigned to patients when they first enter the ED by a triage provider (a physician or nurse), reflecting the patient's medical need as perceived by the provider. The scores identify how long the provider believes a patient can wait safely for treatment or admission. Similarly to authors assessing racial and gender wage disparities, we compare the Elixhauser scores to the triage scores to identify any systematic differences in how providers assess need across Black and White patients (Blau & Kahn, 2017; Western & Pettit, 2005).[2]

We observe that the triage scores rise in step with the Elixhauser scores up to a threshold but then plateau as the Elixhauser scores – and associated mortality – begin to climb. Moreover, at each value of the Elixhauser index, providers deem acuity significantly lower for Black than for White patients.[3] This pattern is consistent with racial bias in clinical assessments of severity (Balsa & McGuire, 2003; Hoffman et al., 2016; Sax et al., 2023; Schulman et al., 1999). Figure III provides further evidence: Regardless of whether patients are admitted or discharged from the ED, Black patients are systematically assigned lower triage scores than White patients.

Collectively, the patterns in Figures II and III suggest that (1) the Elixhauser scores can proxy true medical need for both Black and White patients, (2) providers may underestimate actual medical need among patients with higher Elixhauser scores, and (3) providers may systematically underestimate medical need for Black patients.

[2]The Elixhauser indices for the individual patients in our sample are not directly available to providers. That is, they are known to the econometrician, but not the clinician. Providers can observe the diagnosis codes for the comorbidities we use to compute these indices, but the EMR does not summarize them in index form. Providers thus must predict how diagnoses individually and jointly may affect a patient's likelihood of in-hospital mortality–a cognitively demanding task.

[3]The fitted curves for Black and White patients in Figure II, Panel (b), differ significantly whether or not we constrain curvature to be equal. The difference is driven by the intercept (we do not detect differences in the linear and quadratic terms), suggesting a downward shift in triage scores for Black patients across the mortality index.

3 Racial Disparities in Mortality with Hospital Strain

In this section, we present our main finding: As hospital capacity strain increases, Black patients are likelier to die than White patients.

3.1 Research design

We estimate strain-related racial disparities in outcomes using the following two related specifications:

$$Y_{iht} = \sum_{p=1}^{10} \eta_p \, 1\{\text{Strain}_{ht} = p\} + \sum_{p=1}^{10} \beta_p \left(\text{Race}_i \times 1\{\text{Strain}_{ht} = p\}\right) + \boldsymbol{X}_i'\boldsymbol{\theta} + \delta_j + \pi_{ht} + \varepsilon_{iht} \tag{1a}$$

$$Y_{iht} = \gamma \, \text{HighStrain}_{ht} + \phi \, \text{Race}_i + \beta \big(\text{HighStrain}_{ht} \times \text{Race}_i\big) + \boldsymbol{X}_i'\boldsymbol{\theta} + \delta_j + \pi_{ht} + \varepsilon_{iht} \tag{1b}$$

In the first specification, we regress admission i's in-hospital mortality on ten indicators representing the strain faced by hospital h at patient arrival hour t and an interaction between each indicator with patient race, Race_i, equal to 1 if the patient is Black. The coefficients β_p on the interaction terms provide the marginal effect of patient race on mortality at each strain level. We use this specification in the figures depicting mortality patterns for Black and White patients across the ten deciles.

In the second specification, we regress in-hospital mortality on high hospital strain, HighStrain_{ht}, defined as the 10th decile, with low strain corresponding to the pooled 1st–9th deciles; patient race (Race_i); and the interaction of race and high strain. The coefficient β on this interaction recovers how mortality changes from low to high strain for Black relative to White patients. We use this specification for hypothesis testing and robustness checks.

Identification in our model requires the assumption that, conditional on covariates, no omitted variables are correlated with [Race $\times$ HighStrain]. Accordingly, in Equations 1a and 1b, we include a rich set of patient-level covariates and fixed effects ($\boldsymbol{X}_i$). We adjust for patient demographics by including age, age squared, and patient sex. To account for differences in care patterns for uninsured and low-socioeconomic-status patients (Doyle, 2005), we include patient insurance status. To adjust for differences in patient acuity or medical need, we add fixed effects for Elixhauser mortality and readmission index scores and fixed effects for the sum of the comorbidities used to construct these scores. Fixed effects for the attending physician of record (δ_j) control for differences across the clinical service lines represented in our sample (e.g., surgery) and physician-specific differences in clinical practice patterns, to the extent that the attending physician, who is typically assigned at discharge, provided care during critical periods affecting patient outcomes. Finally, hospital $\times$ year $\times$ month $\times$ day of week $\times$ hour of day fixed effects and hospital $\times$ major holiday fixed effects – a total of

9,910 time-related fixed effects – to account for typical patient flows over these dimensions. These ensure our measure of capacity strain is net of these potentially expected averages and therefore less likely to reflect anticipated strain. As an example of the type of variation these time-related fixed effects exploit, we are able to compare, for instance, differences in mortality between Black and White patients arriving at Hospital 1 in 2016 between 10 and 11 AM across the four Tuesdays during the month of August.

3.2 Main mortality result

Figure IV Panels (a) and (b) plot, respectively, the β_p coefficients denoting the marginal effects of race on mortality from Equation 1a and the fitted values of mortality at each decile. The regression coefficients from Equation 1b are presented in Table 2 Column 1.

Together, the results demonstrate a relative increase in mortality for Black, but not White, patients as the hospital reaches maximum capacity. Figure IV shows that while racial differences in mortality increase most sharply at decile 10, the differences start emerging around decile 7. The interaction coefficient in Table 2 Column 1 indicates that as strain rises to the highest decile, the probability of in-hospital mortality, relative to the pooled outcomes at deciles 1–9 of strain, increases by 0.53 pp more for Black than for White patients (SE = 0.23 pp, p = 0.02). Thus, when hospitals face their highest decile of strain, Black patients are 0.7 pp likelier to die than White patients ($p < 0.001$). However, at lower strain (deciles 1–9, pooled), the Black–White mortality gap is just 0.2 pp (p = 0.054). In short, high strain more than triples the racial mortality gap.

3.2.1 Specification checks

Our substantive findings are robust to specification changes and inclusion of additional covariates. We present coefficients from different specifications ranging from parsimonious to more saturated in Table 2 Columns 2–9. First, we present estimates from the more parsimonious versions, where we include only the hospital × year fixed effects and no other covariates (Column 2) and then a logistic regression version of this parsimonious model (Column 3); our estimates remain unchanged. Next, we present estimates from models with increasingly richer sets of covariates. We add 722 diagnosis-related group (DRG) fixed effects to Equation 1a to ensure that we are making comparisons within a specific health condition (note that the DRGs are identified ex post and so may be endogenous; Column 4). We interact hospital strain with demographic covariates in X_i to account for their potentially differential effects on mortality by strain (e.g., older patients may have greater mortality at high strain than younger patients; Column 5). We interact patient race with the demographic covariates in X_i for similar reasons (e.g., older Black patients may have higher mortality than White patients;

Column 6). To address concerns that our using weighted, summative indices of comorbidities such as the Elixhauser score instead of adjusting for the individual index components may make the estimates more prone to bias (Möller, Bliddal, & Rubin, 2021), we estimate models with separate indicators for each of the 31 comorbidities comprising the Elixhauser indices (instead of fixed effects for the weighted mortality and readmission indices; Column 7). In Column 8, we reestimate our main model with standard errors clustered at the level of the admission date. In Column 9, we add covariates for five measures of abnormal vital signs (heart rate, temperature, blood pressure, respiratory rate, oxygen saturation) first recorded during the inpatient encounter.[4] The results remain substantively similar in magnitude and statistical significant across these specifications.

We next show that our results are robust to alternative definitions of "high strain." Our main specification uses the 90th percentile (decile 10) as the cutoff. Here, we reestimate Equation 1a defining high strain at thresholds from the 50th to the 95th within-hospital percentiles of filled beds and low strain as the pooled percentiles below each cutoff. Our mortality estimates remain robust across these alternative definitions (Figure A.4, Panel (a)). The patterns align with those shown in Figure IV, where the Black–White mortality gap begins to widen around decile 7. Moreover, our study uses data from only two hospitals, so external validity is an important concern. To partially address this, we show that the increase in racial mortality gaps with strain can be qualitatively observed within each hospital separately (Table A.2). While the interaction between race and high strain is statistically significant at the 10% level for Hospital 1 but not Hospital 2, we cannot reject equality of the interaction coefficients across hospitals ($p = 0.35$).

Finally, we estimate Equation 1b with hospital × arrival date × hour fixed effects. This specification compares mortality outcomes between Black and White patients who arrive at the same hospital in the same date × hour cell, and therefore experience the same level of capacity strain. Because the hospital × date × hour cells are small – on average, only 2.48 patients arrive at a given hospital in a given hour, and patients of both races are present in only 47% of those hours – including 26,629 date × hour fixed effects leaves little variation for studying mortality, which is itself a rare outcome. This variation is further reduced when we add the full set of covariates in our main specification (e.g., 667 physician fixed effects and 230 Elixhauser-related fixed effects, among others).

[4] "Abnormal" measures on each vital sign are designated as follows: (i) body temperature above 37.8 C (100 F) or below 35 C (95 F), (ii) diastolic blood pressure below 60 or systolic blood pressure below 90, (iii) respiratory rate below 12 or above 20, (iv) heart rate below 60 or above 100, and (v) oxygen saturation below 90. The thresholds denoting abnormal vitals reflect common criteria used by clinicians. We exclude these "first recorded" vital signs from our main specification because they may be endogenous. For example, if high strain leads to delayed care for Black patients, their vitals might be measured later than those of White patients – and that delay could itself cause differences in the readings. Moreover, vital signs such as oxygen saturation and respiratory rate are not recorded uniformly for all patients and may reflect provider assessments, which themselves could be biased. In this way, the initial vital values may partly reflect differences in care allocation rather than true baseline condition.

For this reason, we report estimates from hospital × date × hour models both with the "full" covariate set and with a "sparse" specification including only race, high strain, and their interaction. Across both models, the point estimates are very similar to those from our main specification, with larger standard errors as expected given the loss of degrees of freedom (Figure A.5, Panel (a)).

3.3 Heterogeneity in mortality by medical need

We next investigate heterogeneity by patient medical need. We categorize patients as high or low medical need by whether their Elixhauser score falls above or below the sample median. We find that our mortality results are driven by high-need patients: The coefficient on the interaction of patient race with high strain and high medical need is 0.0081 ($p = 0.09$) (Table A.3, Column 1). This finding is also apparent in Figure V, where Panels (a) and (b) plot the fitted values of mortality for low- and high-need patients. While on average the latter are likelier to die than the former at all levels of strain, the mortality difference between Black and White patients widens under high strain for high-need patients but remains stable for low-need patients. High-need patients experience significantly greater *absolute* increases in mortality risk either because their medical needs are inherently more challenging to diagnose, especially under conditions of strain, or because they are disproportionately vulnerable to potential reductions in care quantity and quality under increasing strain.

We further examine whether the racial mortality gap changes *relatively* more with strain. That is, we assess percentage rather than absolute changes using a Poisson model for high- and low-need patients (we drop the physician fixed effects to make this model tractable). The triple interaction coefficient is negative and statistically nonsignificant: Among low-need patients, moving from lower to high strain is associated with roughly a 76% higher mortality rate for Black patients relative to White patients, whereas among high-need patients the corresponding increase is about 33% (see Table A.3, Column 2).[5]

4 Ruling out Patient Selection as a Mechanism

The mortality increases we find for Black but not White patients with increasing hospital strain could be due to differential selection into or out of the inpatient sample. If increased strain results in the marginal admission being a sicker Black patient or the marginal discharge being a healthier White patient, then comparable mortality patterns could arise even absent any race-based rationing of care. These possibilities challenge our core identification assumption, which requires that conditional

[5] The triple interaction coefficient is positive when we examine absolute, i.e., *percentage-point*, changes but is negative when we examine relative, i.e., *percent*, changes because baseline mortality differs substantially between low- and high-need patients. The former have a much lower baseline mortality rate (0.3%) than the latter (3.4%), so even small absolute increases in mortality among low-need patients translate into large percentage increases. This makes the effect of strain in percentages appear more pronounced for low-need than for high-need patients.

on covariates, no omitted variables are correlated with [Race × HighStrain]. Importantly, strain need not be completely random (a condition unlikely to hold), and Black and White patients need not be comparable (which they certainly are not). We conduct several tests to rule out this patient selection mechanism and bolster our identification assumption.

The evidence presented thus far is already inconsistent with differential selection as an explanation for our mortality findings. For instance, Black and White patients arrive to the hospital at similar times and experience comparable hospital strain on arrival (Figure I). Our results remain substantively unchanged after we account for detailed time-related fixed effects. While in our main specification, we include hospital × year × month × day of week × hour of day fixed effects, we also confirm robustness under an even more stringent specification that includes hospital × date × hour fixed effects, which should absorb virtually all time-related variation in hospital strain that might influence patient selection. Further, providers themselves consistently perceive Black patients – both admitted and discharged – to be less acutely ill than White patients despite their higher observed mortality (Figure III), which makes it unlikely that selective admission of sicker Black patients drives our findings. Last, the operations and health services literature highlights how hard it is for hospitals to anticipate strain, particularly hour-to-hour changes, and, even when they do anticipate it, to mitigate its adverse effects by adjusting staffing or patient admission patterns (Arogyaswamy et al., 2021; Hoe, 2022; Song et al., 2020). We add to this evidence against patient selection with tests to more conclusively rule out selection as the primary mechanism driving our mortality findings.

4.1 Testing for selective admission

We evaluate differential admission into the inpatient sample by race and strain with a series of balance tests (Pei, Pischke, & Schwandt, 2019). We estimate Equation 1b with a range of observed patient characteristics as the outcome variables: age, sex, insurance status, number of comorbidities, Elixhauser mortality index, and Elixhauser readmission index. Each regression includes all the right-hand side terms from Equation 1b, excluding only the outcome of interest. Across these tests, we find no evidence of significantly differential associations between patient race and strain (Table A.4).

However, tests focusing on individual patient attributes may be underpowered, which would obscure our ability to detect meaningful selection. We address this concern in two ways. First, we create a summed index of each of these covariates, standardized and coded such that higher values imply greater mortality risk, and use this index as our outcome of interest in Equation 1b, omitting the patient covariates used in the index. The estimates from this regression in Table 3 Column 2 suggest that even jointly, these outcomes do not change differentially by race with increasing strain.[6]

[6]Note that when we re-estimate Equation 1b excluding the patient covariates used to construct this summed index,

Second, we estimate a race-specific predicted mortality measure using logistic regressions of mortality on the same covariates included in the summed index. This predicted measure closely tracks observed mortality in our sample (Figure A.3) and offers two advantages over the summed index: (i) It places predicted mortality on the same probability scale as actual mortality, and (ii) it assigns data-driven weights to the covariates, giving more influence to stronger predictors of mortality. We then use this predicted mortality as the outcome in Equation 1b and find no race-specific changes in mortality with increasing strain (Table 3, Column 3). Notably, the 95% confidence interval for the joint effect of [Race x High Strain] on *predicted* mortality excludes the point estimate for the joint effect of [Race x High Strain] on *observed* mortality, suggesting that selection on observables is unlikely to explain the main result.

Then, we examine the possibility of selective admission of certain patient types: patients coming in through the ED and those arriving to the ED by ground or air ambulance. While it is unclear ex ante whether these patients are higher or lower need than other patients, differential shifts in their shares by race may still reveal differential changes in the composition of Black versus White patients with increasing strain. Thus, in Table 3, we test for differential changes in the ED sample in Black and White patients' probability of hospital admission by strain (Column 8) or of arrival to the ED via ambulance (Column 9). We also assess in the full inpatient sample whether Black patients are likelier to be admitted from the ED as strain increases (Column 4). We find no evidence of changes in patient composition along these dimensions with strain.

Finally, we check for selection into the sample by evaluating the possibility of differential patient admission by the patient's presenting complaint. We classify these complaints across five primary "topic" themes, which we identify from the free-text field in the EMR that captures each patient's *reason for admission*. This field is filled in manually by health care providers as soon as a patient arrives at the hospital during triage. We use an ML technique, Latent Dirichlet Allocation (LDA; Blei, Ng, & Jordan, 2003), to identify these themes. LDA is a Bayesian topic-modeling technique that treats each admission note as a combination of a fixed number of latent topics (here, we specify five), each defined by characteristic word distributions. It then estimates, for every patient note, the probability that it belongs to each of these five themes, yielding quantitative measures of presenting complaints that we use to test for differential selection by race and hospital strain. (Appendix A provides additional details.)

We estimate Equation 1b with each of these five themes as outcomes. The *distribution* of themes identified in the *reason for admission* text field does not change differentially with strain across Black

the coefficient on Black $\times$ High Strain is 0.0056 (SE $= 0.0024$, $p < 0.05$).

and White patients (see Table A.5). We take this as additional evidence that patient selection with capacity strain is unlikely a threat to our identification.

4.2 Testing for selective discharge

Like selective admission, differential discharge of patients by race and hospital strain could account for the strain-related mortality patterns we observe. For instance, if, during periods of high strain, White patients are likelier to be discharged to hospice, where death is imminent by definition, this could bias our estimates, especially given that hospice referral rates are known to be higher for White than for Black patients (Asch et al., 2021). Selective transfer of patients to other hospitals under strain is another potential explanation for the strain-related mortality we observe. Relatedly, if sicker White patients arriving at times of high strain are likelier to be discharged prematurely, we may see evidence of their having higher 30-day readmission rates. We find no evidence of differential discharge to hospice (Table 3 Column 5), transfer to other hospitals (Column 6), or probability of 30-day readmission (Column 7) by strain. This suggests that selective discharge is unlikely to account for our main mortality results.

5 Evidence for Race-Based Rationing as a Mechanism

Having found no evidence that patient selection drives our mortality findings, we now present results consistent with race-based rationing of scarce resources being a key mechanism for our mortality findings. Health care resources can be rationed along many dimensions, most of them not observable in EMRs. We focus on two that are observable in our data: wait times and provider effort, recognizing that evidence of rationing by race on these margins may imply similar rationing on other, unmeasured aspects of care.[7] We also consider whether providers assess Black patients' needs less accurately than White patients' under capacity strain, though we view this analysis as speculative.

5.1 Rationing of Inpatient Beds

The first margin of rationing we examine is wait time for an inpatient bed, given its importance in health care and the broader literature on the economics of rationing (Barzel, 1974; Lindsay & Feigenbaum, 1984). Intuitively, patients with more urgent needs should be prioritized for beds, but in practice, wait times do not not always align with patient acuity (Chan & Gruber, 2020). Black

[7] We also investigate rationing along other dimensions of care quality/intensity: (i) ICU admission, (ii) inpatient stay length, and (iii) inpatient charges. These results are presented in Table A.6. How we should interpret these results, however, is unclear: These dimensions capture care much further downstream from the time of arrival, i.e., when the patient is exposed to a given level of strain in our analysis. Thus, these inputs can be interpreted as both resources that may have been rationed and as *outcomes* of rationing decisions themselves. For example, Black patients appear more likely to be admitted to the ICU when strain reaches its highest decile; however, because ICU admission often precedes death—and Black patients in our sample also face higher mortality—this pattern likely mirrors care correlated with mortality rather than distinct allocation behavior. Consequently, we consider these tests less informative for our primary question but report them for completeness.

patients, in particular, may wait longer for care than White patients (Lu, Hanchate, & Paasche-Orlow, 2021), a finding that holds for public services other than health care, as well (Chen et al., 2022; Holt & Vinopal, 2023).

Our timestamped EMR data allow us to calculate wait times from arrival at the hospital to placement in an inpatient bed down to the second. We can compute wait times for 86% of our main sample, specifically for those patients arriving at the hospital units where waiting for the next step of care typically occurs (Main Registration, ED, Surgical Arrival Services, and Observation).[8]

5.1.1 Results

Figure VI shows the marginal effects of race on wait times (Panel (a)) and the fitted values of wait times at each strain decile (Panel (b)), both estimated from Equation 1a. The regression estimates from Equation 1b appear in Table 4 Column 1. Black patients wait longer than White patients across all levels of strain, and their wait times rise more sharply than White patients' as strain increases. On average, Black patients wait 0.95 hours longer for an inpatient bed than White patients at lower strain but 1.3 hours longer – a 37% relative increase –than White patients at high strain. Indeed, we see a "bump" in wait times at the highest decile of hospital strain (Figure VI Panel (a)), mirroring the pattern for mortality (Figure IV Panel (a)). This finding also helps address remaining concerns that the widening racial mortality gap under strain may be driven by differences in medical need between Black and White patients that are unobserved to the researcher. If Black patients were, in reality, sicker in ways we do not observe, we should expect them to receive care faster, not to wait longer.

Specification checks: We conduct the same set of specification checks for our wait time results as for our mortality results. First, in Table 4 Columns 2–8, we assess sensitivity to model sparsity in specifications ranging from one with only hospital × year fixed effects (Column 2) to more saturated models with DRG fixed effects, individual Elixhauser comorbidities, vital signs, and other controls. Second, we again test alternative percentile cutoffs at which we define high strain. Specifically, we define high strain as occupancy at or above thresholds ranging from the 50th to the 95th percentiles (Figure A.4 Panel (b)). Third, we estimate models with hospital × arrival date × hour fixed effects with both the full and sparse specifications (Figure A.5 Panel (b)). Across all the specifications, our results remain both statistically and substantively consistent.

Heterogeneity by patient medical need: Next, we assess whether the race-based allocation of inpatient beds by wait time occurs in addition to, or in place of, allocation by patient medical need. We reestimate our main Equations 1a and 1b but now interact the race and strain variables with patient

[8] Our mortality findings remain robust within the subsample of patients with calculated wait times (see Table A.9).

need. Recall that high- and low- need patients are identified by Elixhauser mortality index values above and below the sample median, respectively. The triple interaction term is not statistically significant (Table A.8), but we highlight three key patterns from Figure VI Panel (b), which plots the fitted wait times for Black and White patients by clinical need. First, while strain increases wait times for all patients, we find evidence that medical need becomes less predictive of bed allocation as strain increases: While low-need patients, regardless of race, wait longer for care than high-need patients at lower capacity levels, this gap vanishes at higher strain.

Second, at almost every capacity decile (except decile 2), the difference in wait times between Black and White patients is larger than the difference between low- and high-need patients. Moreover, as strain increases, though the difference in wait times between high- and low-need patient disappears, it is this large difference in wait times between Black and White patients that persists. In other words, if we could observe only patient wait times at high strain, we would be able to distinguish Black and White patients but not high- and low-need ones.

Third, high-need Black patients wait longer than low-need White patients at almost every capacity decile (except decile 2), with this difference being the largest at the highest decile.[9] The persistence of this racial disparity even at the lowest decile of strain suggests that the wait-time differences cannot be solely attributed to resource scarcity. The disparities could instead reflect underlying taste-based or statistical discrimination at the provider or institutional level that exists independently of resource constraints but becomes more pronounced as those constraints increasingly bind. However, because our data are not suited for testing for omitted variables correlated with strain × race × medical need, we cannot confirm this mechanism.

5.2 Rationing of Provider Effort

The second margin of rationing we examine is provider effort. Effort is required for providers to accurately assess clinical need and appropriately diagnose and treat patients. However, like time, effort is a scarce resource: Under personnel or resource constraints, providers may wish to ration their effort in favor of sicker, higher-risk patients, but as with wait times, there may be suboptimal rationing of effort by need. That effort is rationed by race is suggested by findings from the literature: For example, systematic racial biases held by providers manifested into lower intensities of effort exerted for Black patients in experimental studies using simulated clinical cases (Stepanikova, 2012), and surveys of physicians' implicit attitudes reveal greater bias against Black patients under cognitive stressors (Johnson et al., 2016).

[9] Our main specification accounts for patient insurance status. This finding holds after we additionally adjust for mode of arrival to the ED.

In contrast to wait times, provider effort is not directly measured even in rich clinical data like ours. Proxies can be found for specific types of clinical encounters (Chan, 2018), but effort is difficult to capture for the entire course of a hospitalization and when patients come in with many different health conditions and care needs. To circumvent these challenges, we analyze unstructured medical notes. We focus on the *reason for admission*, a text field manually filled out at the time of the patient's hospital arrival (typically as a triage step) by a physician or nurse.[10] This note is the first note recorded in the medical record and serves primarily as a transcription or detailed summary of the patient's stated reason for visiting the ED or for being admitted. Documentation in this field is consequential because it triggers a cascade of downstream clinical decisions – such as diagnostic priority, specialist involvement, nursing resources, and bed assignments – that shape care quality and affect patient outcomes. For example, an admission reason specifying respiratory distress might lead providers to immediately prioritize diagnostic tests such as pulse oximetry, respiratory therapy, and a monitored bed. A less precise entry, such as "weakness," can route the patient to lower-acuity care areas, delaying diagnosis and intervention. However, because analysis of textual data – particularly clinical text – is a nascent research area with limited methodological guidance, we interpret these analyses as suggestive about mechanisms.

We infer effort (at least that of the provider who documents the reason for admission) by creating a summed index of five (standardized) features of the documented text: (i) character counts, (ii) average word length, (iii) subjectivity, (iv) polarity, and (v) number of adjectives. We code each such that higher values imply more effort. We discuss our choice and construction of each below and then present results from estimating Equations 1a and 1b using the summed index and each component as outcomes. Descriptive statistics of each of the five components comprising the provider effort index by race and decile of hospital strain are provided in Figure A.9.

The first and second features, respectively, are the *character count*, the number of characters used in the text field, and *average word length*, the total character count divided by number of words used.[11] We use these measures in our index as greater character counts and word lengths in free-text fields have been shown to be strongly correlated with effort in surveys and laboratory experiments (Galesic & Bosnjak, 2009; Tausczik & Pennebaker, 2010; Yadav, Prabhu, & Chandy, 2007).

The third and fourth features of the text field that are included in this index are the note's *subjectivity score* and *polarity score*.[12] To compute these scores, we analyze the sentiment in the

[10] As for wait times, a minority (27%) of our main sample has an empty *reason for admission* field. Reestimating Equation 1b on this subsample with mortality as the outcome does not change our main result (Table A.9).

[11] We use a median split of the character count because the variable is so skewed: Among the nonmissing text data, the skew is 4.5. The skew of the average word length, for comparison, is 1.6.

[12] Beyond measuring effort, subjectivity and polarity can predict mortality risk (Waudby-Smith et al., 2018).

medical documentation using the TextBlob library (see Appendix B for details). Subjectivity refers to the degree of opinion vs. factual information encoded in a text on a 0–1 scale, with higher scores implying greater subjectivity. For example, a statement such "Patient describes pain as a 3 out of 10" could receive a subjectivity score of zero. In contrast, a statement such as "I am deeply optimistic that this patient may survive" or "Patient describing unimaginable levels of pain" could receive a subjectivity score of 0.9. Greater subjectivity leaves room for implicit biases in documentation and care based on that documentation (Bloche, 2001; Schrader & Lewis, 2013). Further, subjective writing is known to be linguistically simpler (Ehret & Taboada, 2021) and requires less effort to produce than objective content (Best, Floyd, & McNamara, 2008). Thus, higher values of the subjectivity score may imply lower levels of effort (we reverse-code this score, such that increasing values imply greater effort, in constructing our index).

Polarity refers to the emotional leaning of the text, i.e., whether there is an opinion expressed in the text and whether the opinion is negative or positive. Polarity scores range from -1 to 1 (with -1, 0, and 1 representing negative, neutral, and positive text). A statement such as "Patient lives nearby" could receive a polarity score of 0; "Severe chest pain worsening overnight" could receive a polarity score of -1; and "Symptoms improving significantly since yesterday" could reflect a polarity score of +1. Polarity may affect effort because accurately documenting a sicker patient's condition typically requires highlighting negative or concerning clinical details clearly and urgently. Thus, more positive polarity scores might reflect less careful or less urgent documentation, potentially leading providers to underestimate patient acuity. We thus predict that higher levels of polarity may imply lower levels of effort (and thus reverse code this score, as well, when constructing our index).

The fifth feature included in our provider effort index is the *number of adjectives*, which measures the presence of common suffixes in the text.[13] Adjectives provide nuance, context, and depth about the object of description, especially in medicine – e.g., "a single rapidly-growing red-to-purplish small pedunculated raspberry-like lesion" (Roland, 1967) – and, more importantly, can express changes (for example, in the condition of a patient's illness) along a scalar dimension (Kennedy & Levin, 2008). Greater use of adjectives is consistent with greater effort.

5.2.1 Results

Using the summed provider effort index as the outcome in Equations 1a and 1b, we present evidence consistent with race-based rationing of effort under increasing strain in Figure VII and Table 5.

[13] We use this method as opposed to relying on a part-of-speech tagger in TextBlob, which would perform poorly on medical text because it uses both definition and context in a sentence to identify adjectives and the *reason for admission* field does not always follow regular English grammatical and syntactical structure.

As with wait times, we observe that average provider effort is significantly lower for Black than for White patients. At lower levels of strain (pooled deciles 1–9), Black patients receive 0.10 units lower provider effort than White patients. However, this Black–White disparity doubles in magnitude to 0.21 units at the highest decile of strain (an increase of 0.11 pts, or 0.04 standard deviations; see Table 5 Column 1). Figure VII Panel (a) also reveals a distinct "dip" in provider effort at decile 10, mirroring the "bump" seen at high strain for both mortality and wait times. (Results for each component of the index are reported separately in Table A.10 and visualized in Figure A.10).

Thus, we find evidence of rationing not only of material resources (inpatient beds) but also cognitive resources. Cognitive resource allocation can affect patient outcomes in broader, subtler, and potentially more harmful ways than material resources alone, influencing diagnostic accuracy, treatment decisions, and both objective and subjective aspects of care quality in ways that material constraints may not fully capture. While audits can readily detect biases in the allocation of beds, medications, or even time, it is significantly more challenging to detect and quantify rationing of cognitive resources such as provider effort and attention or related resources such as respect, empathy, and trust.

Specification checks: We repeat the same specification checks for our provider effort results as for mortality and wait times. The magnitude of our estimates generally holds across the different covariate specifications (Table 5 Columns 2–8), definitions of high strain at alternate percentile cutoffs (Figure A.4 Panel (c)), and the inclusion of arrival date × hour fixed effects (Figure A.5 Panel (c)). That some coefficients are more imprecise likely reflects that our measure of effort, derived from several features assessed from a single text field in the medical record, is noisier than the more objective measures of wait time and mortality.

Heterogeneity by patient medical need: Parallel to our analysis of wait times, we examine differential changes in effort by strain, race, and medical need (Figure VII Panel (b)). The triple-interaction coefficient is not statistically significant, and the estimates are noisier than those for wait times; however, we highlight two descriptive patterns in the figure. First, on average, medical need determines provider effort allocations for Black but not White patients (i.e., high-need Black patients receive more effort than low-need Black patients, while White high- and low-need patients receive similar effort). Second, as with wait times, patient race becomes increasingly predictive of provider effort allocations as strain intensifies. At the lowest strain decile, effort does not differ significantly by race, but by the tenth decile, the racial disparity is pronounced. Indeed, at the highest decile of strain, *high-need* Black patients receive less provider effort than *low-need* White patients (Figure VII Panel (b)).

5.3 Assessment of medical need by providers under strain

Thus far, we have found that the widening racial mortality gap is concentrated among high-need patients and that, under rising strain, the allocation of beds and provider effort appears to become decoupled from clinical need. One possibility is that strain diminishes providers' ability or willingness to assess mortality risk accurately more for Black than for White patients, leading to greater reliance, whether conscious or unconscious, on race in resource allocation. We assess this possibility by examining heterogeneity in mortality by the objectivity of patient symptoms, recognizing that data limitations make this exercise less conclusive.

With rising strain, providers' reliance on race instead of medical need should be less pronounced in cases where medical need is more objectively measurable or where symptoms clearly indicate the appropriate course of action. Thus, we focus on patients presenting with *chest pain* and/or *shortness of breath* – common, high-risk complaints that may signal myocardial infarction, pulmonary embolism, or acute asthma and that typically trigger a protocolized work-up, including chest radiography, electrocardiography, and cardiac enzyme testing (Kasper et al., 2015). (These are the only objective symptom categories we can cleanly isolate in our data.)

We extract mentions of these two symptoms by searching the free-text *reason for admission* field for relevant phrases, including variants such as "sob" for "shortness of breath." We observe no strain-related racial mortality disparities for these more objective symptoms; however, our main mortality pattern remains intact for the subsample of less protocolized, more subjective symptoms (Figure VIII Panels (c) and (d)). We interpret these results as suggestive. Although the triple interaction between patient race, strain, and symptom objectivity is not statistically significant (Table A.3, Column 3), the estimates are imprecise, likely reflecting limited power given the rarity of mortality and the fact that only 9.1% of admissions involve chest pain or shortness of breath.[14]

5.4 Discussion

Our findings are consistent with multiple theories of discrimination, and our data do not allow us to disentangle them. We briefly discuss each in turn, noting where they overlap and diverge.

Taste-based discrimination reflects provider animus or dispreference toward Black patients. Such preferences may be nonbinding when capacity is ample, but as trade-offs become unavoidable, they may manifest in reduced effort and longer wait times for Black patients (Balsa & McGuire, 2003).

Statistical discrimination is distinct in that it does not require animus. Instead, providers use race as a signal for unobserved patient characteristics when making clinical decisions under uncertainty.

[14]We also examine whether there are systematic strain-induced differences in provider assessments of medical need (i.e., triage scores) by patient race, but are underpowered to detect such differences (see Appendix C).

Under high strain, when time and information are scarce, providers may lean more heavily on group-level beliefs – for instance, assuming that Black patients will adhere less to treatment and therefore expecting smaller returns to investing scarce effort, even when a patient's true clinical needs are identical to those of a White patient. Worse outcomes for Black patients may then appear to "confirm" these priors, amplifying disparities over time.

Implicit bias is closely related to statistical discrimination but emphasizes the cognitive mechanism. Laboratory evidence shows that when cognitive or material resources become scarce, decision-making shifts toward fast, heuristic ("Type I") processing that amplifies reliance on stereotypes (Bodenhausen, 1990; Kahneman, 2011).[15] In clinical settings, such heuristics may distort assessments of risk or assumptions about treatment adherence, systematically disadvantaging Black patients (Balsa & McGuire, 2003; Schulman et al., 1999). The key distinction from standard models of statistical discrimination is the emphasis on automatic, nondeliberative processing rather than Bayesian updating. In practice, however, the two channels are difficult to separate, as both operate through group-level beliefs and both intensify under resource constraints.

Finally, systemic discrimination offers a distinct channel that does not require any contemporaneous provider bias (Bohren, Hull, & Imas, 2022). Black patients may arrive at the hospital with missing diagnostic information due to prior underdiagnosis or with elevated disease severity from prior undertreatment in outpatient settings (Darity, 2022). Providers under strain may then fail to account for these accumulated disadvantages, inadvertently reinforcing existing disparities. This mechanism is consistent with our finding that racial gaps in wait times and effort widen with hospital strain.

6 Conclusion

Our study documents important connections between resource scarcity, discrimination, and rationing. Focusing on the high-stakes setting of hospital care, we find that resource scarcity arising from high levels of capacity strain unmasks discrimination, resulting in race-based rationing of care and worse health outcomes for Black patients. We first show that Black–White disparities in patient death – an extreme consequence of rationing on the basis of race rather than need – increase markedly at the highest levels of hospital capacity strain. We are unable to reconcile this pattern with racial differences in the types of patients admitted or discharged at high versus low levels of strain. Instead, we provide evidence of rationing by race in wait times and provider effort as potential mechanisms, showing that racial disparities in wait times and our measures of provider effort exist at all levels of

[15] For example, under extreme time pressure, participants were more likely to misidentify harmless tools as guns after being primed with Black faces – a bias significantly reduced with more deliberation time (Payne, 2001). Similarly, inducing a scarcity mindset lowered thresholds for categorizing racially ambiguous faces as Black and reduced resource allocations to these faces (Krosch & Amodio, 2014).

strain but become particularly large at the highest levels of strain.

Our findings inform future work seeking to identify discrimination in health care and a variety of real-world settings. Specifically, they demonstrate how systemic shocks – even ones lasting only hours or days – can throw into sharper relief discrimination that is otherwise difficult to discern without external manipulation (such as in audit studies). In this sense, our research echoes the work in health care by Gandhi (2020), who shows that dynamic changes in bed availability can help identify whether nursing homes cherry-pick patients on the basis of their health insurance. Our approach and findings echo work from other contexts documenting how adverse economic shocks led to increases in lynchings of Black individuals in the US (Beck & Tolnay, 1990) and in civil conflict globally (Miguel, Satyanath, & Sergenti, 2004) and how improved economic prospects (due to either labor demand shocks or policy-led increases in access to new markets) helped reduce intergroup inequality (Aizer et al., 2020; Black & Strahan, 2001). By leveraging such shocks, we can gain insights into mechanisms of discrimination otherwise hidden by the routine operations of markets and institutions.

References

Aizer, A., Boone, R., Lleras-Muney, A., & Vogel, J. (2020). *Discrimination and racial disparities in labor market outcomes: Evidence from wwii* (tech. rep.). National Bureau of Economic Research.

Akbarpour, M., Budish, E., Dworczak, P., & Kominers, S. D. (2024). An economic framework for vaccine prioritization. *The Quarterly Journal of Economics*, *139*(1), 359–417.

Alsan, M., Garrick, O., & Graziani, G. (2019). Does Diversity Matter for Health? Experimental Evidence from Oakland. *American Economic Review*, *109*(12), 4071–4111. https://doi.org/10.1257/aer.20181446

Alsan, M., & Wanamaker, M. (2018). Tuskegee and the health of black men. *The quarterly journal of economics*, *133*(1), 407–455.

Anderson, D. M., Crost, B., & Rees, D. I. (2020). Do economic downturns fuel racial animus? *Journal of Economic Behavior & Organization*, *175*, 9–18.

André, C., & Platteau, J.-P. (1998). Land relations under unbearable stress: Rwanda caught in the malthusian trap. *Journal of Economic Behavior & Organization*, *34*(1), 1–47.

Antunes, R. d. A., Gonçalves, E. d. S., Bernardino, L. G., Casalecchi, J. G. S., Grebot, I. B. d. F., & de Moraes Jr, R. (2023). Influence of Economic Scarcity on Race Perception [Publisher: SAGE Publications Sage CA: Los Angeles, CA]. *Psychological Reports*, 00332941231169666.

Arogyaswamy, S., Vukovic, N., Angela Keniston, A., Apgar, S., Bowden, K., Kantor, M., Diaz, M., McBeth, L., & Burden, M. (2021). The Impact of Hospital Capacity Strain: A Qualitative Analysis of Experience and Solutions at 13 Academic Medical Centers. *Journal of General Internal Medicine*, *epub before print*.

Arrow, K. J. (1998). What Has Economics to Say about Racial Discrimination? *The Journal of Economic Perspectives*, *12*(2), 91–100. Retrieved March 24, 2025, from https://www.jstor.org/stable/2646963

Asch, D. A., Islam, M. N., Sheils, N. E., Chen, Y., Doshi, J. A., Buresh, J., & Werner, R. M. (2021). Patient and Hospital Factors Associated With Differences in Mortality Rates Among Black and White US Medicare Beneficiaries Hospitalized With COVID-19 Infection. *JAMA network open*, *4*(6), e2112842. https://doi.org/10.1001/jamanetworkopen.2021.12842

Auer, D. (2022). Firing discrimination: Selective labor market responses of firms during the covid-19 economic crisis. *PloS one*, *17*(1), e0262337.

Baert, S., Cockx, B., Gheyle, N., & Vandamme, C. (2015). Is there less discrimination in occupations where recruitment is difficult? *ILR Review*, *68*(3), 467–500.

Balsa, A. I., & McGuire, T. G. (2003). Prejudice, clinical uncertainty and stereotyping as sources of health disparities. *Journal of health economics*, *22*(1), 89–116.

Barzel, Y. (1974). A theory of rationing by waiting. *The Journal of Law and Economics*, *17*(1), 73–95.

Bayer, P., Ferreira, F., & Ross, S. L. (2018). What drives racial and ethnic differences in high-cost mortgages? the role of high-risk lenders. *The Review of Financial Studies*, *31*(1), 175–205.

Beck, E. M., & Tolnay, S. E. (1990). The killing fields of the deep south: The market for cotton and the lynching of blacks, 1882-1930. *American Sociological Review*, 526–539.

Berkebile-Weinberg, M. M., Krosch, A. R., & Amodio, D. M. (2022). Economic scarcity increases racial stereotyping in beliefs and face representation. *Journal of Experimental Social Psychology*, *102*, 104354.

Best, R. M., Floyd, R. G., & McNamara, D. S. (2008). Differential competencies contributing to children's comprehension of narrative and expository texts [Place: United Kingdom]. *Reading Psychology*, *29*(2), 137–164. https://doi.org/10.1080/02702710801963951

Black, S. E., & Strahan, P. E. (2001). The division of spoils: Rent-sharing and discrimination in a regulated industry. *American Economic Review*, *91*(4), 814–831.

Blau, F. D., & Kahn, L. M. (2017). The Gender Wage Gap: Extent, Trends, and Explanations. *Journal of Economic Literature*, *55*(3), 789–865. https://doi.org/10.1257/jel.20160995

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, *3*(Jan), 993–1022.

Bloche, M. G. (2001). Race and discretion in American medicine. *Yale J. Health Pol'y L. & Ethics*, *1*, 95.

Bodenhausen, G. V. (1990). Stereotypes as Judgmental Heuristics: Evidence of Circadian Variations in Discrimination. *Psychological Science*, *1*(5), 319–322. Retrieved April 17, 2025, from https://www.jstor.org/stable/40062735

Bohren, J. A., Hull, P., & Imas, A. (2022). *Systemic discrimination: Theory and measurement* (tech. rep.). National Bureau of Economic Research.

Breza, E., Kaur, S., & Shamdasani, Y. (2021). Labor rationing. *American Economic Review*, *111*(10), 3184–3224.

Burke, L. G., Frakt, A. B., Khullar, D., Orav, E. J., & Jha, A. K. (2017). Association between teaching status and mortality in us hospitals. *Jama*, *317*(20), 2105–2113.

Card, D., & Krueger, A. B. (1996). School resources and student outcomes: An overview of the literature and new evidence from north and south carolina. *Journal of economic Perspectives*, *10*(4), 31–50.

Chan, D. C. (2018). The efficiency of slacking off: Evidence from the emergency department. *Econometrica*, *86*(3), 997–1030.

Chan, D. C., & Gruber, J. (2020). Provider discretion and variation in resource allocation: The case of triage decisions. *AEA Papers and Proceedings*, *110*, 279–283.

Chandra, A., & Skinner, J. S. (2003). Geography and racial health disparities. National bureau of economic research Cambridge, Mass., USA.

Charles, K. K., & Hurst, E. (2002). The transition to home ownership and the black-white wealth gap. *Review of Economics and Statistics*, *84*(2), 281–297.

Chen, M. K., Haggag, K., Pope, D. G., & Rohla, R. (2022). Racial disparities in voting wait times: Evidence from smartphone data. *Review of Economics and Statistics*, *104*(6), 1341–1350.

Cook, B. L., McGuire, T. G., & Zaslavsky, A. M. (2012). Measuring racial/ethnic disparities in health care: Methods and practical issues. *Health services research*, *47*(3pt2), 1232–1254.

Darity, W. A. (2022). Position and Possessions: Stratification Economics and Intergroup Inequality. *Journal of Economic Literature*, *60*(2), 400–426.

Darity, W. A., & Mason, P. L. (1998). Evidence on Discrimination in Employment: Codes of Color, Codes of Gender. *Journal of Economic Perspectives*, *12*(2), 63–90. https://doi.org/10.1257/jep.12.2.63

Doyle, J. J. (2005). Health insurance, treatment and outcomes: Using auto accidents as health shocks. *Review of Economics and Statistics*, *87*(2), 256–270.

Dube, O., MacArthur, S. J., & Shah, A. K. (2023). *A cognitive view of policing* (tech. rep.). National Bureau of Economic Research.

Ehret, K., & Taboada, M. (2021). The interplay of complexity and subjectivity in opinionated discourse. *Discourse Studies*, *23*(2), 141–165. https://doi.org/10.1177/1461445620966923

Eli, S. J., Logan, T. D., & Miloucheva, B. (2023). The Enduring Effects of Racial Discrimination on Income and Health. *Journal of Economic Literature*, *61*(3), 924–940.

Elixhauser, A., Steiner, C., Harris, D. R., & Coffey, R. M. (1998). Comorbidity measures for use with administrative data. *Medical care*, 8–27.

Galesic, M., & Bosnjak, M. (2009). Effects of questionnaire length on participation and indicators of response quality in a web survey. *Public opinion quarterly*, *73*(2), 349–360.

Gandhi, A. (2020). Picking your patients: Selective admissions in the nursing home industry. *Available at SSRN*, *3613950*.

Glaeser, E., & Gyourko, J. (2018). The economic implications of housing supply. *Journal of economic perspectives*, *32*(1), 3–30.

Glaeser, E. L., & Luttmer, E. F. P. (2003). The Misallocation of Housing Under Rent Control. *American Economic Review*, *93*(4), 1027–1046. https://doi.org/10.1257/000282803769206188

Goddu, A. P., O'Conor, K. J., Lanzkron, S., Saheed, M. O., Saha, S., Peek, M. E., Haywood, C., & Beach, M. C. (2018). Do words matter? stigmatizing language and the transmission of bias in the medical record. *Journal of general internal medicine*, *33*, 685–691.

Green, R. K., & White, M. J. (1997). Measuring the benefits of homeowning: Effects on children. *Journal of urban economics*, *41*(3), 441–461.

Haushofer, J., Lowes, S., Musau, A., Ndetei, D., Nunn, N., Poll, M., & Qian, N. (2023). Stress, Ethnicity, and Prosocial Behavior. *Journal of Political Economy Microeconomics*, *1*(2), 225–269. https://doi.org/10.1086/722367

Hoe, T. P. (2022). Does Hospital Crowding Matter? Evidence from Trauma and Orthopedics in England. *American Economic Journal: Economic Policy*, *14*(2), 231–62.

Hoffman, K. M., Trawalter, S., Axt, J. R., & Oliver, M. N. (2016). Racial bias in pain assessment and treatment recommendations, and false beliefs about biological differences between blacks and whites. *Proceedings of the National Academy of Sciences*, *113*(16), 4296–4301.

Holt, S. B., & Vinopal, K. (2023). Examining inequality in the time cost of waiting. *Nature Human Behaviour*, 1–11.

Institute of Medicine. (2003). *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care* (B. D. Smedley, A. Y. Stith, & A. R. Nelson, Eds.). National Academies Press (US). Retrieved September 24, 2021, from http://www.ncbi.nlm.nih.gov/books/NBK220358/

Jackson, C. K., Wigger, C., & Xiong, H. (2021). Do school spending cuts matter? evidence from the great recession. *American Economic Journal: Economic Policy*, *13*(2), 304–335.

Johnson, T. J., Hickey, R. W., Switzer, G. E., Miller, E., Winger, D. G., Nguyen, M., Saladino, R. A., & Hausmann, L. R. M. (2016). The Impact of Cognitive Stressors in the Emergency Department on Physician Implicit Racial Bias. *Academic emergency medicine*, *23*(3), 297–305. https://doi.org/10.1111/acem.12901

Jost, T. S. (2022). Considering race and ethnicity in covid risk assessments—legal concerns and possible solutions. *New England Journal of Medicine*, *387*(6), 481–483.

Kadri, S. S., Sun, J., Lawandi, A., Strich, J. R., Busch, L. M., Keller, M., Babiker, A., Yek, C., Malik, S., Krack, J., et al. (2021). Association between caseload surge and covid-19 survival in 558 us hospitals, march to august 2020. *Annals of Internal Medicine*, *174*(9), 1240–1251.

Kahneman, D. (2011). *Thinking, fast and slow*. macmillan.

Kasper, D., Fauci, A., Hauser, S., Longo, D., Jameson, J., & Loscalzo, J. (2015). *Harrison's principles of internal medicine, 19e* (Vol. 1). Mcgraw-hill New York, NY, USA:

Kennedy, C., & Levin, B. (2008). Measure of change: The adjectival core of degree achievements.

King, J. J. C., Powell-Jackson, T., Hargreaves, J., Makungu, C., & Goodman, C. (2023). Does increased provider effort improve quality of care? evidence from a standardised patient study on correct and unnecessary treatment. *BMC health services research*, *23*(1), 190.

Klein, B., Ogbunugafor, C. B., Schafer, B. J., Bhadricha, Z., Kori, P., Sheldon, J., Kaza, N., Sharma, A., Wang, E. A., Eliassi-Rad, T., et al. (2023). Covid-19 amplified racial disparities in the us criminal legal system. *Nature*, 1–7.

Kocher, K. E., Dimick, J. B., & Nallamothu, B. K. (2013). Changes in the source of unscheduled hospitalizations in the united states. *Medical care*, *51*(8), 689–698.

Kohn, R., Harhay, M. O., Weissman, G. E., Anesi, G. L., Bayes, B., Greysen, S. R., Ratcliffe, S. J., Halpern, S. D., & Kerlin, M. P. (2019). Ward Capacity Strain: A Novel Predictor of Delays in Intensive Care Unit Survivor Throughput. *Annals of the American Thoracic Society*, *16*(3), 387–390. https://doi.org/10.1513/AnnalsATS.201809-621RL

Krosch, A. R., & Amodio, D. M. (2014). Economic scarcity alters the perception of race. *Proceedings of the National Academy of Sciences*, *111*(25), 9079–9084.

Krosch, A. R., Tyler, T. R., & Amodio, D. M. (2017). Race and recession: Effects of economic scarcity on racial discrimination. *Journal of personality and social psychology*, *113*(6), 892.

Lindsay, C. M., & Feigenbaum, B. (1984). Rationing by waiting lists. *The American economic review*, *74*(3), 404–417.

Lu, F. Q., Hanchate, A. D., & Paasche-Orlow, M. K. (2021). Racial/ethnic disparities in emergency department wait times in the united states, 2013–2017. *The American Journal of Emergency Medicine*, *47*, 138–144.

Martin, S., & Smith, P. C. (1999). Rationing by waiting lists: An empirical investigation. *Journal of Public Economics*, *71*(1), 141–164.

Miguel, E., Satyanath, S., & Sergenti, E. (2004). Economic shocks and civil conflict: An instrumental variables approach. *Journal of political Economy*, *112*(4), 725–753.

Möller, S., Bliddal, M., & Rubin, K. H. (2021). Methodical considerations on adjusting for charlson comorbidity index in epidemiological studies. *European Journal of Epidemiology*, *36*(11), 1123–1128.

Moore, B. J., White, S., Washington, R., Coenen, N., & Elixhauser, A. (2017). Identifying increased risk of readmission and in-hospital mortality using hospital administrative data. *Medical care*, *55*(7), 698–705.

Neprash, H. T., Everhart, A., McAlpine, D., Smith, L. B., Sheridan, B., & Cross, D. A. (2021). Measuring primary care exam length using electronic health record data. *Medical care*, *59*(1), 62–66.

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science (New York, N.Y.)*, *366*(6464), 447–453. https://doi.org/10.1126/science.aax2342

Payne, B. K. (2001). Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon [Place: US]. *Journal of Personality and Social Psychology*, *81*(2), 181–192. https://doi.org/10.1037/0022-3514.81.2.181

Pei, Z., Pischke, J.-S., & Schwandt, H. (2019). Poorly measured confounders are more useful on the left than on the right. *Journal of Business & Economic Statistics*, *37*(2), 205–216.

Price-Haywood, E. G., Burton, J., Fort, D., & Seoane, L. (2020). Hospitalization and mortality among black patients and white patients with covid-19. *New England Journal of Medicine*, *382*(26), 2534–2543.

Roland, C. G. (1967). Nouns and Adjectives. *JAMA*, *201*(12), 961–962. https://doi.org/10.1001/jama.1967.03130120069019

Rugh, J. S., & Massey, D. S. (2010). Racial segregation and the american foreclosure crisis. *American sociological review*, *75*(5), 629–651.

Sax, D. R., Warton, E. M., Mark, D. G., Vinson, D. R., Kene, M. V., Ballard, D. W., Vitale, T. J., McGaughey, K. R., Beardsley, A., Pines, J. M., et al. (2023). Evaluation of the emergency severity index in us emergency departments for the rate of mistriage. *JAMA Network Open*, *6*(3), e233404–e233404.

Schrader, C. D., & Lewis, L. M. (2013). Racial disparity in emergency department triage. *The Journal of emergency medicine*, *44*(2), 511–518.

Schulman, K. A., Berlin, J. A., Harless, W., Kerner, J. F., Sistrunk, S., Gersh, B. J., Dube, R., Taleghani, C. K., Burke, J. E., Williams, S., et al. (1999). The effect of race and sex on physicians' recommendations for cardiac catheterization. *New England Journal of Medicine*, *340*(8), 618–626.

Schwab, S. D., & Singh, M. (2023). How power shapes behavior: Evidence from physicians.

Sen, A. (1982). *Poverty and famines: An essay on entitlement and deprivation*. Oxford university press.

Song, H., Tucker, A. L., Graue, R., Moravick, S., & Yang, J. J. (2020). Capacity Pooling in Hospitals: The Hidden Consequences of Off-Service Placement. *Management Science*, *66*(9), 3825–3842. https://doi.org/10.1287/mnsc.2019.3395

Stepanikova, I. (2012). Racial-ethnic biases, time pressure, and medical decisions [Publisher: Sage Publications Sage CA: Los Angeles, CA]. *Journal of health and social behavior*, *53*(3), 329–343.

Stiglitz, J. E. (2012). *The price of inequality: How today's divided society endangers our future*. WW Norton & Company.

Stiglitz, J. E., & Weiss, A. (1981). Credit rationing in markets with imperfect information. *The American economic review*, *71*(3), 393–410.

Stoye, G., & Warner, M. (2023). The effects of doctor strikes on patient outcomes: Evidence from the english nhs. *Journal of Economic Behavior & Organization*, *212*, 689–707.

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: Liwc and computerized text analysis methods. *Journal of language and social psychology*, *29*(1), 24–54.

Vohra, A. S., Khullar, D., Kaushal, R., & Schpero, W. L. (2023). Many Intensive Care Units Were Overloaded While Nearby Hospitals Had Excess Capacity During The COVID-19 Pandemic. *Health Aff (Millwood)*, *42*(7), 937–945.

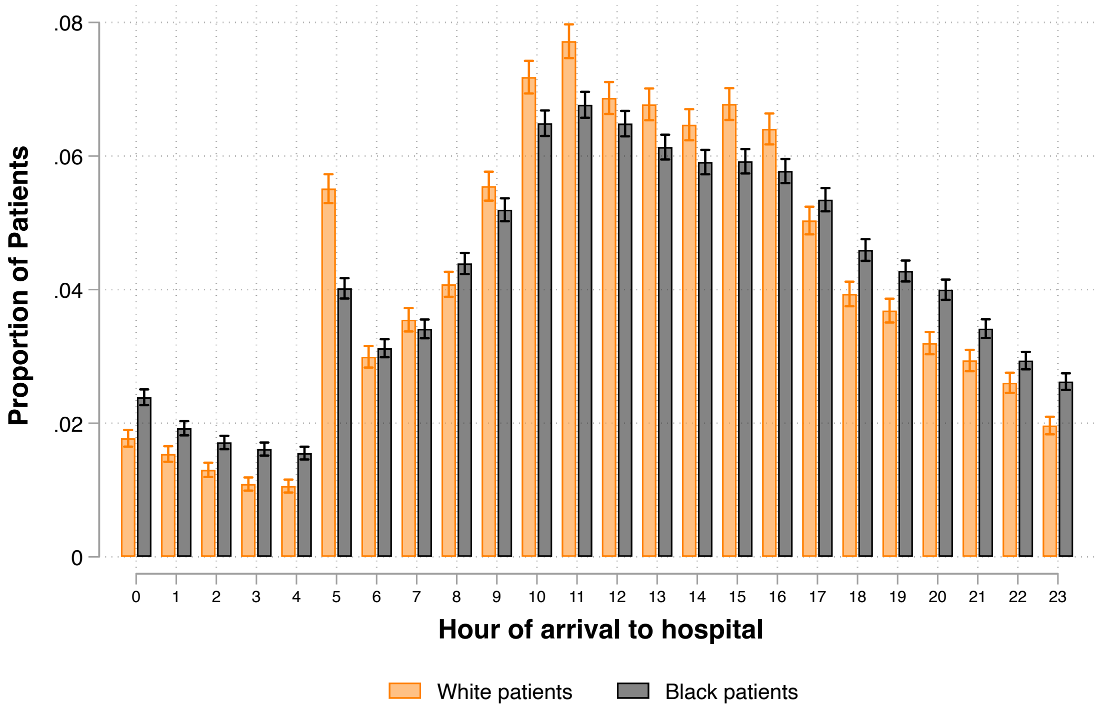
Waudby-Smith, I. E., Tran, N., Dubin, J. A., & Lee, J. (2018). Sentiment in nursing notes as an indicator of out-of-hospital mortality in intensive care patients. *PloS one*, *13*(6), e0198687.

Weitzman, M. L. (1977). Is the price system or rationing more effective in getting a commodity to those who need it most? *The Bell Journal of Economics*, 517–524.

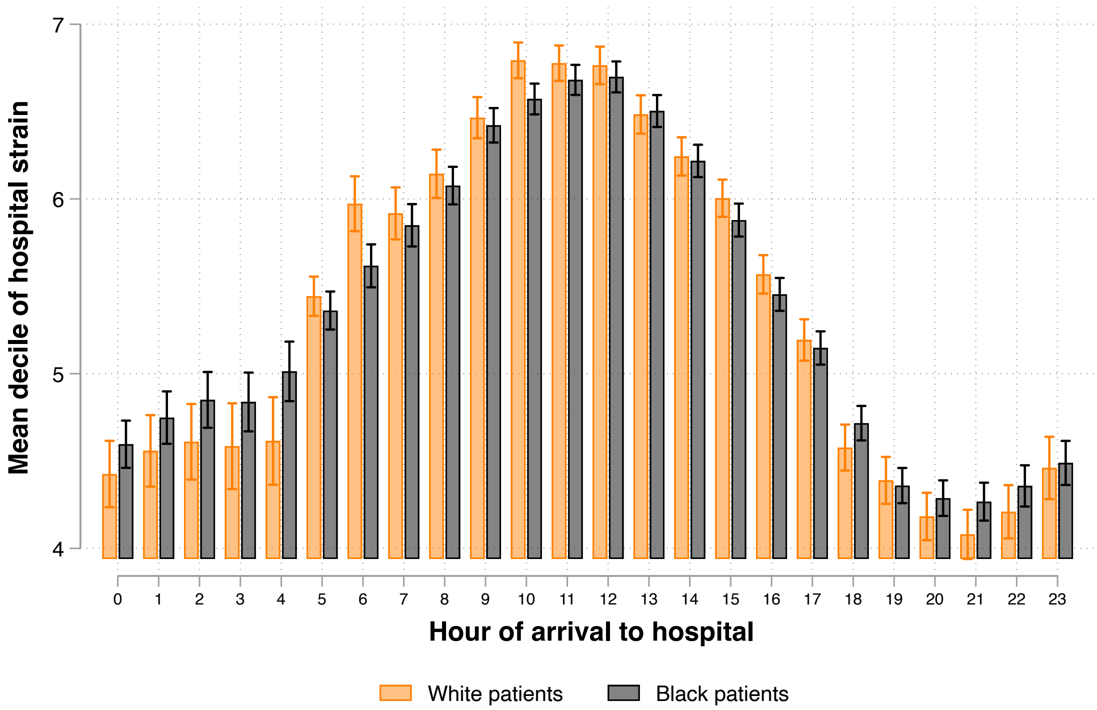
Western, B., & Pettit, B. (2005). Black-White Wage Inequality, Employment Rates, and Incarceration [Place: US]. *American Journal of Sociology*, *111*(2), 553–578. https://doi.org/10.1086/432780

White, D. B., Villarroel, L., & Hick, J. L. (2021). Inequitable access to hospital care—protecting disadvantaged populations during public health emergencies. *New England Journal of Medicine*, *385*(24), 2211–2214.

Yadav, M. S., Prabhu, J. C., & Chandy, R. K. (2007). Managing the future: Ceo attention and innovation outcomes. *Journal of marketing*, *71*(4), 84–101.



Panel (a)



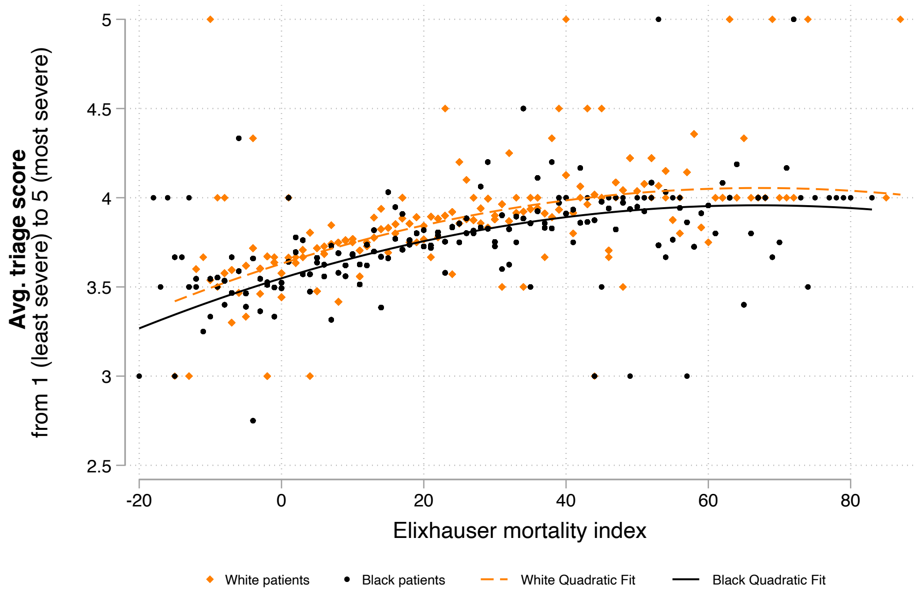
Panel (b)

FIGURE I

Arrival Times and Strain Levels By Race

Note: Panel (a) presents the distribution of White and Black patients by the hour of their arrival to the hospital. Panel (b) presents the mean decile of hospital strain for White and Black patients by the hour of their arrival to the hospital.

Panel (a): Elixhauser mortality index vs. in-hospital mortality (Inpatient sample)

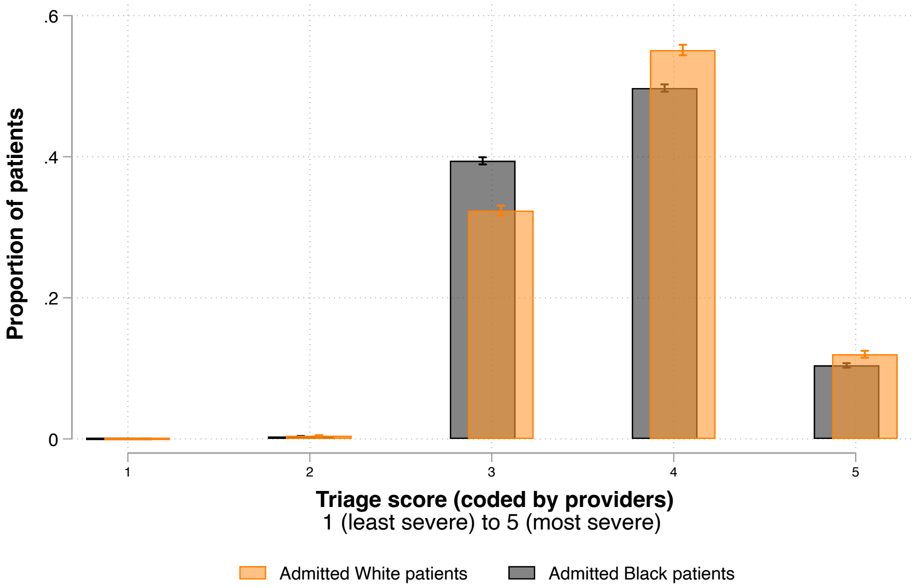


Panel (b): Elixhauser mortality index vs. triage scores (ED sample)

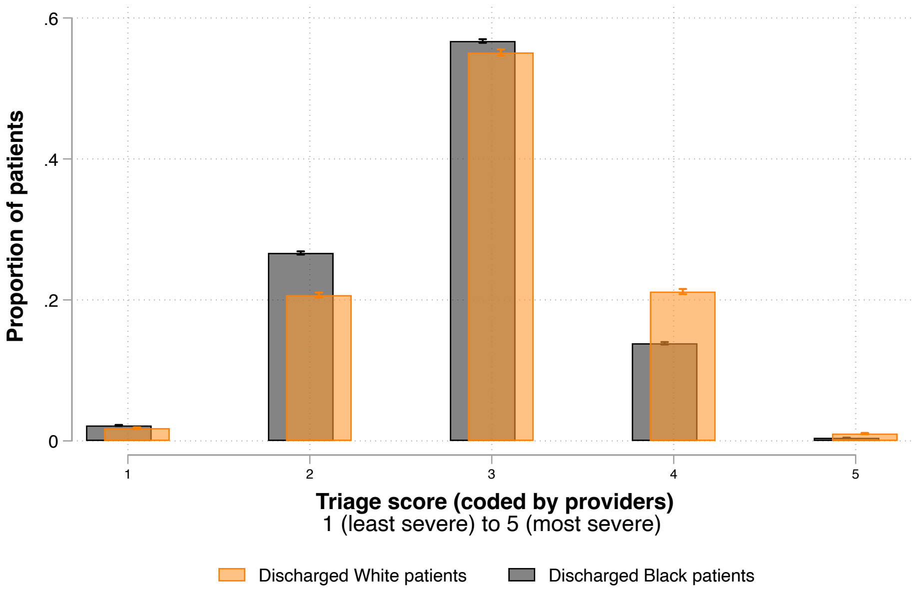
FIGURE II

Elixhauser Mortality Index as a Measure of Patient Need

Note: Panel (a): Distribution of Elixhauser mortality index scores and their (unadjusted) relationship with in-hospital mortality using scatter plots and a (patient-weighted) quadratic best fit line for Black and White patients in the inpatient sample. Panel (b): Unadjusted relationship between Elixhauser mortality index scores and triage scores using scatter plots and a (patient-weighted) quadratic best fit line for Black and White patients in the ED sample.



Panel (a): Triage scores for admitted ED patients

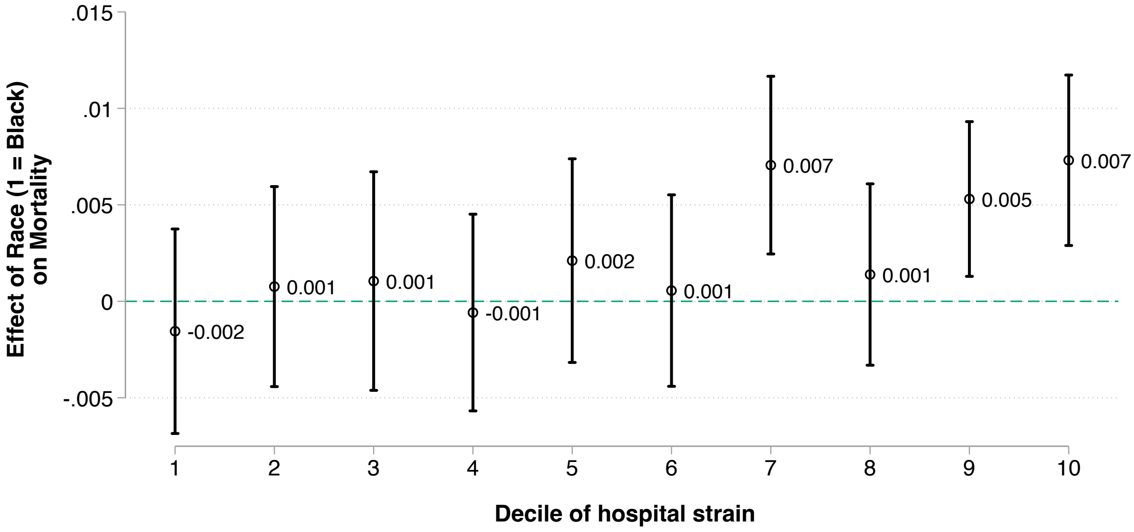


Panel (b): Triage scores for discharged ED patients

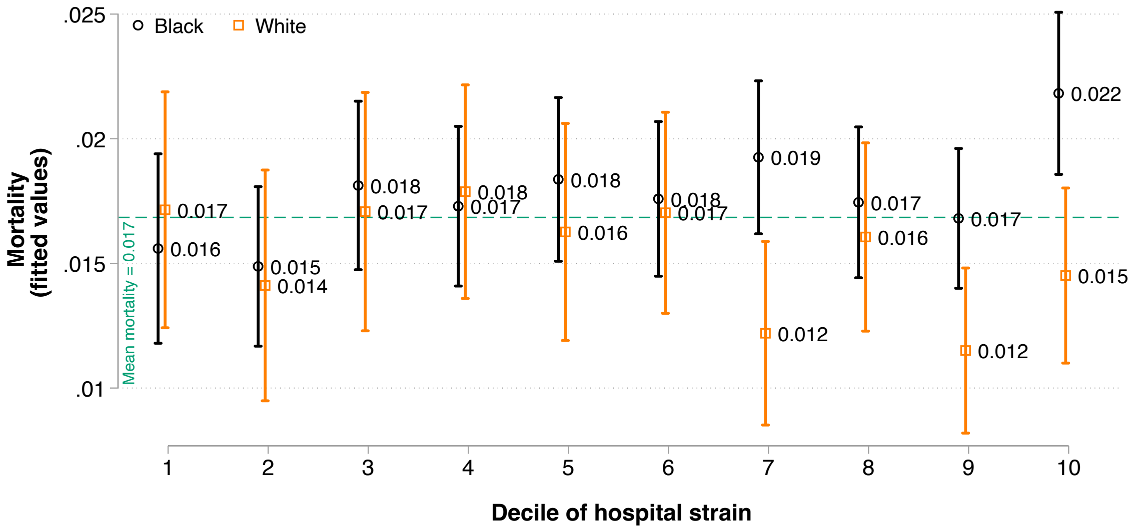
FIGURE III

Triage Scores by Race (ED Sample)

Note: Histogram of triage scores (discrete values of 1–5 , higher triage scores reverse-coded to indicate higher severity) for Black and White patients admitted to the hospital (Panel (a)), and discharged home (Panel (b)), from the ED. Calculated on ED sample only.



Panel (a): Marginal effect of race on in-hospital mortality

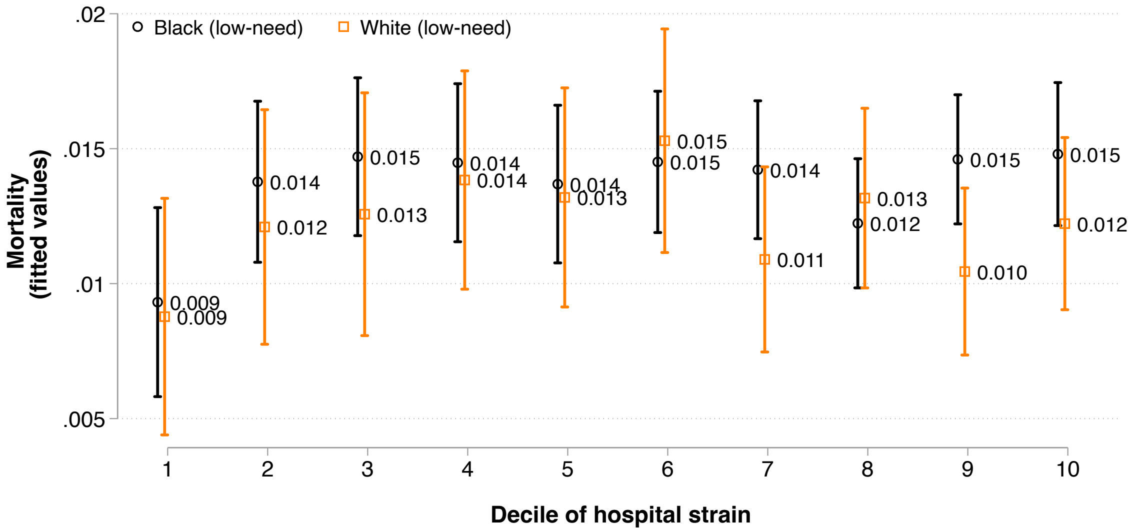


Panel (b): Fitted values of in-hospital mortality

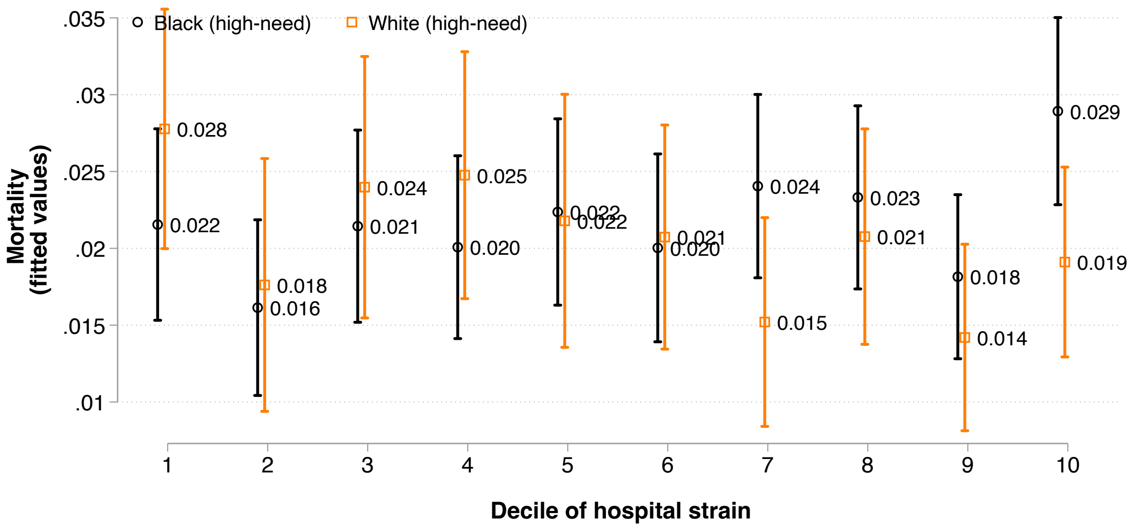
FIGURE IV

Mortality by Race and Strain

Note: Estimates are from Equation 1a, with in-hospital mortality as the outcome. Panel (a) plots the marginal effect of being Black on mortality at each decile of hospital strain at arrival (horizontal dashed line at 0). Panel (b) plots fitted mortality values by race across deciles (horizontal dashed line at sample mean of mortality). All effects are evaluated at the means of other covariates. We show 95% confidence intervals based on robust standard errors. The model includes controls for patient demographics (age, age squared, female), insurance status (insured vs. uninsured), Elixhauser mortality and readmission indices fixed effects, number of comorbidities fixed effects, attending physician fixed effects, hospital × year × month × day of week × hour of day fixed effects, and hospital × major holiday fixed effects. Regression coefficients from Equation 1b are shown in Table 2 Column 1.



Panel (a): Fitted values of in-hospital mortality for low-need patients

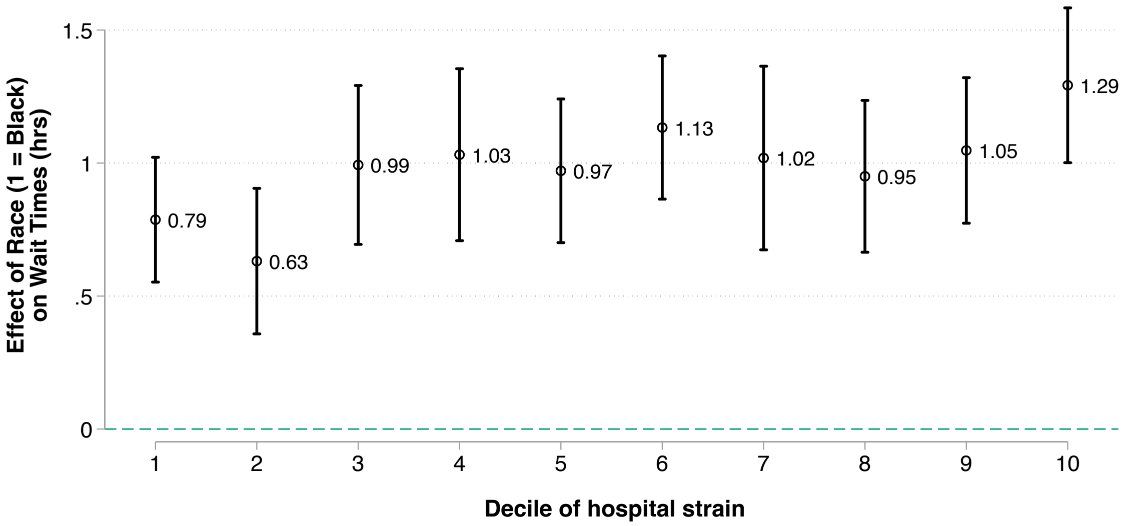


Panel (b): Fitted values of in-hospital mortality for high-need patients

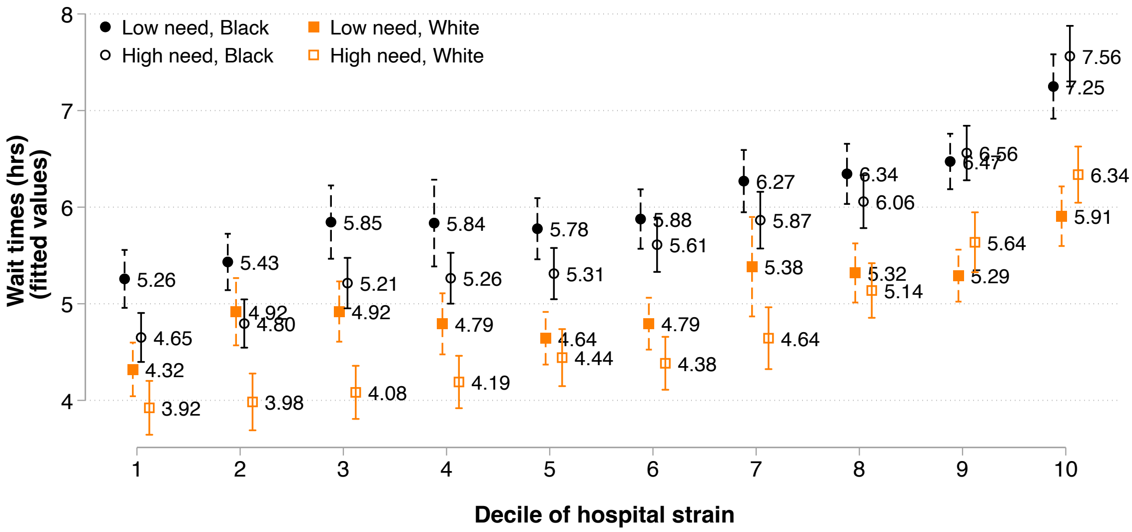
FIGURE V

Heterogeneity in Mortality by Patient Need

Note: Estimates from Equation 1a with in-hospital mortality as the outcome, with three-way interactions between patient race, hospital strain deciles, and patient need. Patient need is defined by Elixhauser mortality index values below (low-need) and above (high-need) the sample median (Panels (a) and (b), respectively). Panels show fitted values of mortality for Black and White patients at each strain decile, evaluated at the means of other covariates, with 95% confidence intervals based on robust standard errors. Regression coefficients are presented in tabular form in Table A.3.



Panel (a): Marginal effect of race on wait times

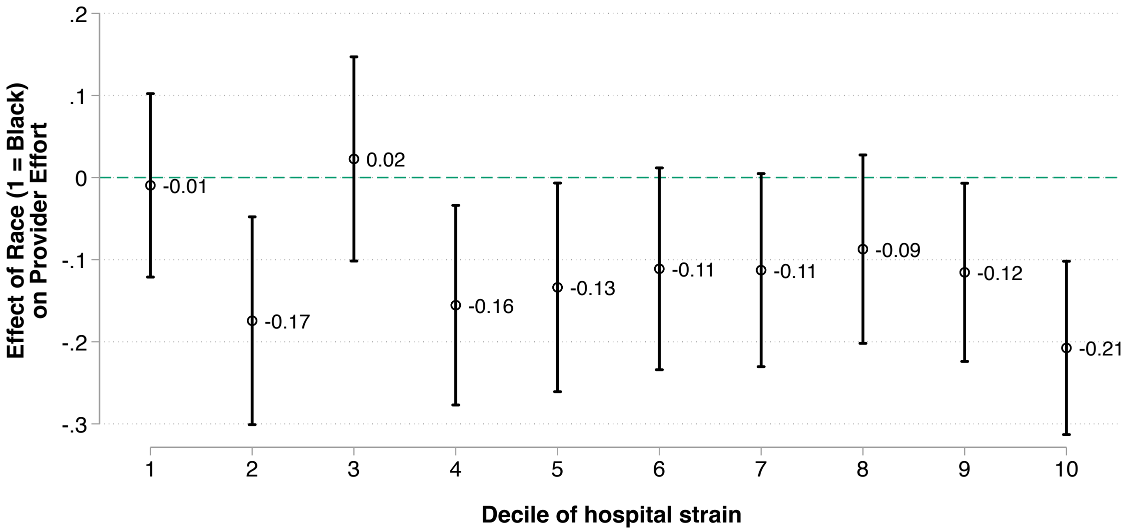


Panel (b): Fitted values of wait times

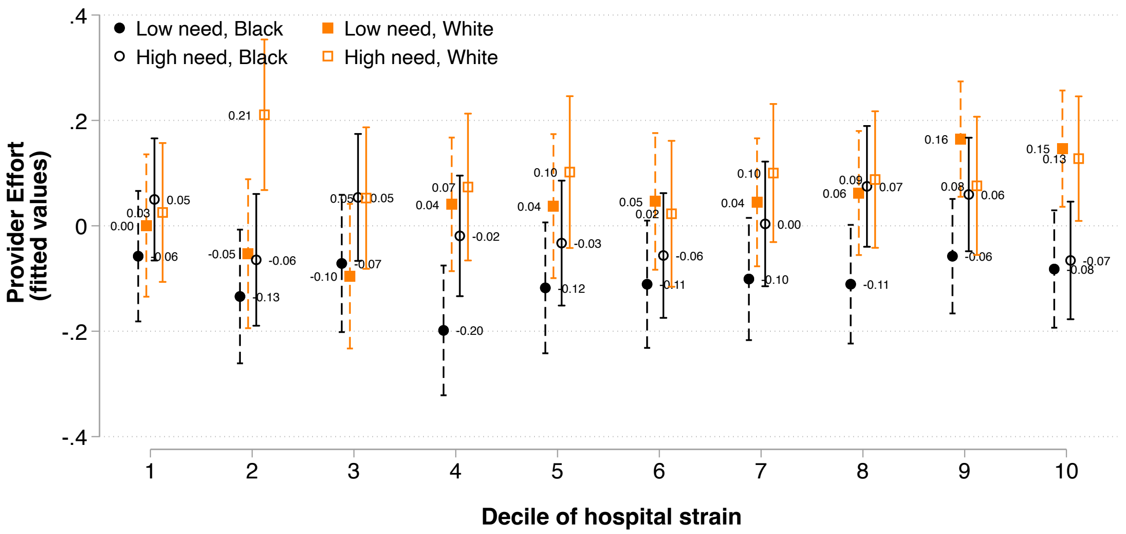
FIGURE VI

Wait Times: Rationing by Race and Medical Need

Note: Estimates are from Equation 1a with wait times (in hours) as the outcome. Panel (a) presents the marginal effect of patient race (Black = 1) at each decile of hospital strain at arrival. Panel (b) plots the fitted values of wait times by race and patient need at each decile of hospital strain (low and high patient need captured by below- and above-median Elixhauser mortality index scores, respectively). All effects and fitted values are evaluated at the means of other covariates, with 95% confidence intervals based on robust standard errors. Regression coefficients from Equation 1b are presented in Table 4.



Panel (a): Marginal effect of race on provider effort

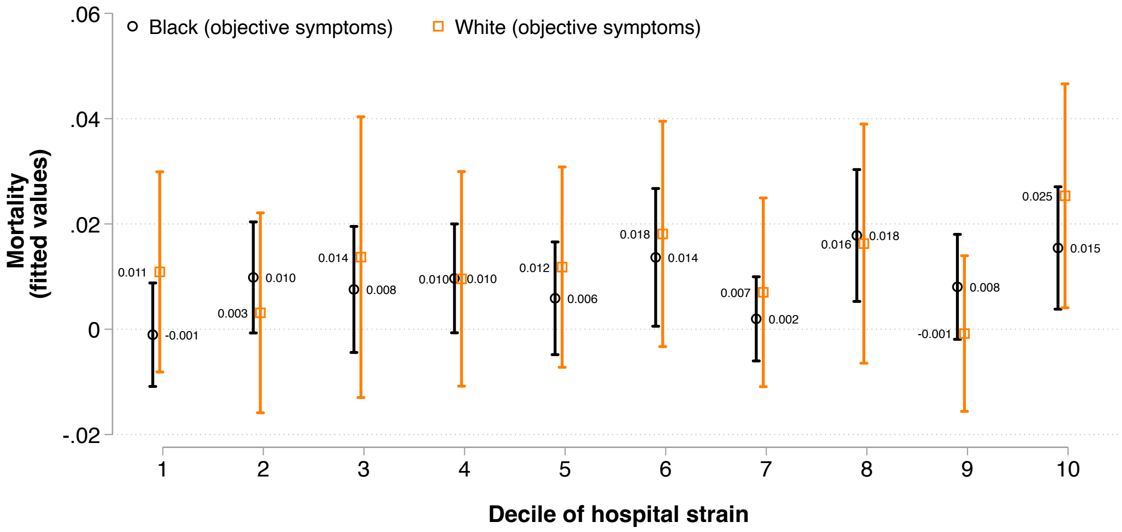


Panel (b): Fitted values of provider effort

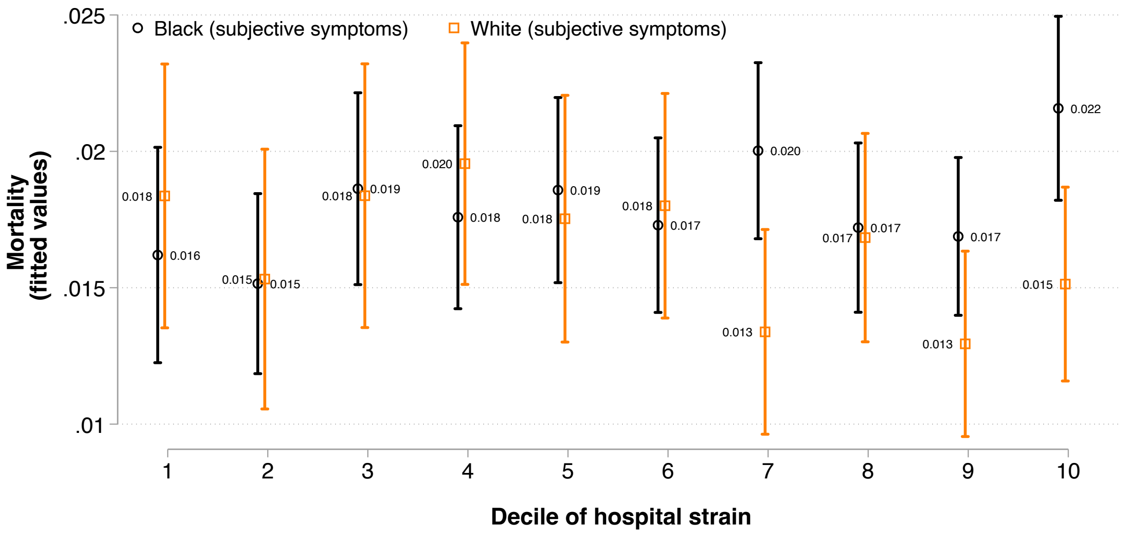
FIGURE VII

Provider Effort: Rationing by Race and Medical Need

Note: Estimates are from Equation 1a with the provider effort index as the outcome. The index is created by summing five standardized features of the *reason for admission* text field (see Figure A.9). Panel (a) presents the marginal effect of patient race (Black = 1) at each decile of hospital strain at arrival. Panel (b) plots the fitted values of the provider effort index by race and patient need at each decile of hospital strain (low and high patient need captured by below- and above-median Elixhauser mortality index scores, respectively). All effects and fitted values are evaluated at the means of other covariates, with 95% confidence intervals based on robust standard errors. Regression coefficients from Equation 1b are presented in Table 5.



Panel (a): Fitted values of in-hospital mortality for patients with more objective symptoms



Panel (b): Fitted values of in-hospital mortality for patients with more subjective symptoms

FIGURE VIII

Heterogeneity in Mortality by Symptom Type (Objective vs Subjective)

Estimates from a version of Equation 1a with in-hospital mortality as the outcome, with three-way interactions between patient race, hospital strain deciles, and symptom type: more objective (Panel (a)) vs. more subjective (Panel (b)). Symptom type is identified from the *reason for admission* text field, with more objective symptoms being those that include mentions of "shortness of breath" or "chest pain"; more subjective symptoms are all others. Panels show fitted values of mortality for Black and White patients at each strain decile, evaluated at the means of other covariates, with 95% confidence intervals based on robust standard errors. Regression coefficients are presented in tabular form in Table A.3 Column 3.

	White Patients Mean (SD)	Black Patients Mean (SD)
Age (yrs.)	59.03 (18.06)	51.95 (19.51)
Uninsured (%)	10.29 (30.38)	14.07 (34.77)
Female (%)	49.71 (50.00)	64.50 (47.85)
Elixhauser mortality score	12.97 (13.12)	12.32 (13.44)
Elixhauser readmission score	23.32 (19.72)	25.30 (22.13)
# of comorbidities	4.08 (2.73)	4.10 (3.05)
ICU admission (%)	27.73 (44.77)	22.19 (41.55)
In-hospital mortality (%)	1.79 (13.24)	1.62 (12.61)
Wait time (hrs)	4.22 (6.70)	6.40 (7.61)
Length of stay (days)	6.72 (9.40)	6.48 (10.23)
30-day readmission (%)	15.61 (36.30)	16.05 (36.71)
Observations	42949	64275

TABLE 1
Summary Statistics

Note: Descriptive statistics for key covariates and outcomes for Black and White patients.

	Dep Var: In-Hospital Mortality								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Black	0.0019*	-0.0008	-0.0473	0.0020**	0.0019*	.	0.0030**	0.0019**	0.0024**
	(0.0010)	(0.0009)	(0.0535)	(0.0010)	(0.0010)	(.)	(0.0010)	(0.0009)	(0.0010)
High strain (=dec 10)	-0.0005	-0.0035**	-0.2187*	0.0001	.	-0.0005	0.0000	-0.0005	-0.0002
	(0.0019)	(0.0018)	(0.1194)	(0.0018)	(.)	(0.0019)	(0.0019)	(0.0017)	(0.0019)
Black × High strain (=dec 10)	0.0053**	0.0068**	0.4116**	0.0043*	0.0051**	0.0053**	0.0042*	0.0053**	0.0052**
	(0.0023)	(0.0024)	(0.1508)	(0.0023)	(0.0024)	(0.0023)	(0.0024)	(0.0022)	(0.0024)
N	106791	107224	107224	106751	106791	106791	106810	106791	102114
R^2	0.292	0.001		0.332	0.292	0.292	0.288	0.292	0.296
Controls	Main Spec	H-Y FEs	Logistic	DRG	x Strain	x Black	Elix comps	Cluster SEs	Vitals

TABLE 2

In-Hospital Mortality: Specification Tests

Note: Coefficients from regressions of in-hospital mortality on race (=1 if Black) and the top decile of hospital strain, using varying specifications:

Col 1: Baseline model (Eq. 1b).
Col 2: Parsimonious controls: hospital × year fixed effects only
Col 3: Same as Col 2, estimated with logit
Col 4: Col 1 + diagnosis-related-group (DRG) fixed effects
Col 5: Col 1 with demographic covariates interacted with strain
Col 6: Col 1 with demographic covariates interacted with race
Col 7: Col 1 replacing Elixhauser indices with 31 comorbidity indicators
Col 8: Col 1 with standard errors clustered by admission date
Col 9: Col 1 + five indicators for abnormal vital signs (heart rate, blood pressure, temperature, respiratory rate, and oxygen saturation)

Sample mean (SD) of in-hospital mortality: 0.017 (0.128). $p < .001^{***}$ $p < .05^{**}$ $p < 0.1^{*}$

	Inpatient Sample							ED Sample	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	In-Hosp Mort	Summed Mort	Predicted Mort	ED Adm	Hospice	Trans out	30-Readm	Admit to hosp	Ambulance
Black	0.0019*	0.0315	0.0004	0.1065***	-0.0041**	-0.0013**	0.0135***	0.0002	0.0467***
	(0.0010)	(0.0224)	(0.0003)	(0.0030)	(0.0013)	(0.0004)	(0.0029)	(0.0019)	(0.0019)
High strain (=dec 10)	-0.0005	-0.0197	-0.0002	-0.0018	-0.0026	-0.0000	0.0039	-0.0005	-0.0141**
	(0.0019)	(0.0467)	(0.0006)	(0.0059)	(0.0026)	(0.0007)	(0.0059)	(0.0055)	(0.0051)
Black × High strain (=dec 10)	0.0053**	0.0776	0.0009	0.0010	0.0017	-0.0000	-0.0045	-0.0057	0.0038
	(0.0023)	(0.0559)	(0.0007)	(0.0070)	(0.0031)	(0.0008)	(0.0072)	(0.0059)	(0.0057)
N	106791	106810	102188	106791	106791	106791	106791	257421	257421
R^2	0.292	0.497	0.227	0.571	0.226	0.128	0.167	0.240	0.113
Mean of outcome	0.017	-0.000	0.017	0.459	0.029	0.002	0.159	0.201	0.203
SD of outcome	0.128	3.667	0.037	0.498	0.167	0.048	0.365	0.401	0.402
Corr. w mortality (White)	N/A	0.153	0.262	0.000	-0.025	-0.007	-0.057	N/A	N/A
Corr. w mortality (Black)	N/A	0.162	0.294	0.033	-0.021	-0.006	-0.056	N/A	N/A
Patient controls	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Physician controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Hospital×time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

TABLE 3

Is Patient Selection Driving Racial Disparities in Mortality?

Note: Regression coefficients and standard errors in parentheses from Equation 1b using the following outcomes:

Inpatient Sample:
Col 1: In-hospital mortality
Col 2: Summed covariate index
Col 3: Predicted mortality from race-specific logistic models
Col 4: Admitted through the ED
Col 5: Discharged to hospice
Col 6: Transferred to another hospital
Col 7: 30-day readmission

ED sample:
Col 8: Admitted to the hospital
Col 9: Arrival by ambulance (ground or air)

N/A means either not applicable, or the relevant variables were not available. $p < .001^{***}$ $p < .05^{**}$ $p < 0.1^{*}$

	Dep Var: Wait Times (hrs)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	0.95***	2.29***	0.94***	0.95***	.	0.90***	0.95***	0.94***
	(0.05)	(0.05)	(0.05)	(0.05)	(.)	(0.05)	(0.05)	(0.05)
High strain (=dec 10)	1.04***	0.88***	1.07***	.	1.04***	1.04***	1.04***	1.03***
	(0.12)	(0.11)	(0.12)	(.)	(0.12)	(0.12)	(0.13)	(0.12)
Black × High strain (=dec 10)	0.35**	0.28*	0.35**	0.39**	0.36**	0.36**	0.35**	0.36**
	(0.15)	(0.16)	(0.15)	(0.16)	(0.15)	(0.15)	(0.16)	(0.15)
N	91898	92581	91877	91898	91898	91917	91898	91680
R^2	0.293	0.026	0.325	0.293	0.293	0.292	0.293	0.294
Controls	Main Spec	H-Y FEs	DRG	× Strain	× Black	Elix comps	Cluster SEs	Vitals

TABLE 4

Wait Times: Specification Tests

Note: Coefficients from regressions of wait times on race (=1 if Black) and the top decile of hospital strain, using varying specifications:

Col 1: Baseline model (Eq. 1b).
Col 2: Parsimonious controls: hospital × year fixed effects only
Col 3: Col 1 + diagnosis-related-group (DRG) fixed effects
Col 4: Col 1 with all covariates interacted with strain
Col 5: Col 1 with all covariates interacted with race
Col 6: Col 1 replacing Elixhauser indices with 31 comorbidity indicators
Col 7: Col 1 with standard errors clustered by admission date
Col 8: Col 1 + five indicators for abnormal vital signs (heart rate, blood pressure, temperature, respiratory rate, and oxygen saturation)

Sample mean (SD) of wait times: 5.46 (7.31). $p < .001^{***}$ $p < .05^{**}$ $p < 0.1^{*}$

	Dep Var: Provider Effort Index							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.10***	-0.36***	-0.07**	-0.09***	.	-0.07**	-0.10***	-0.09***
	(0.02)	(0.02)	(0.02)	(0.02)	(.)	(0.02)	(0.02)	(0.02)
High strain (=dec 10)	0.06	0.10**	0.04	.	0.06	0.06	0.06	0.06
	(0.05)	(0.04)	(0.05)	(.)	(0.05)	(0.05)	(0.04)	(0.05)
Black × High strain (=dec 10)	-0.11*	-0.04	-0.09	-0.12**	-0.11*	-0.11*	-0.11*	-0.11**
	(0.06)	(0.05)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
N	77582	78468	77560	77582	77582	77605	77582	77415
R^2	0.220	0.019	0.247	0.220	0.220	0.218	0.220	0.220
Controls	Main Spec	H-Y FEs	DRG	x Strain	x Black	Elix comps	Cluster SEs	Vitals

TABLE 5

Provider Effort: Specification Tests

Note: Coefficients from regressions of the provider effort index on race (=1 if Black) and the top decile of hospital strain, using varying specifications:

Col 1: Baseline model (Eq. 1b).
Col 2: Parsimonious controls – hospital × year fixed effects only
Col 3: Col 1 + diagnosis-related-group (DRG) fixed effects
Col 4: Col 1 with demographic covariates interacted with strain
Col 5: Col 1 with demographic covariates interacted with race
Col 6: Col 1 replacing Elixhauser indices with 31 comorbidity indicators
Col 7: Col 1 with standard errors clustered by admission date
Col 8: Col 1 + five indicators for abnormal vital signs (heart rate, blood pressure, temperature, respiratory rate, and oxygen saturation)

Sample mean (SD) of provider effort index: 0.002 (2.49). $p < .001^{***}$ $p < .05^{**}$ $p < 0.1^{*}$

Online Appendix

A Identifying themes in the text data with latent Dirichlet allocation

We first provide exploratory evidence of what the *reason for admission* text field contains, and whether – at first glance – it appears starkly different between Black and White patients. We first create word clouds for Black and White patients (Figure A.7), visualizing the most common words and phrases in this text field. We find that similar words – *surgery*, *preadmission*, *pain* – occur the most frequently for both races, though there are some differences. *Pain*, *chest*, and *sob*, which stands for *shortness of breath*, are more frequent among Black patients.

Black and White patients may be described differently in their reason for admission for many reasons: differences in the actual reason, differences in how they describe their symptoms to the providers, and/or differences how providers interpret those symptoms. If differences in words with increasing strain represent differences in the types of patients being admitted to the hospital, this may represent a threat to our identification strategy. We try to test this formally using a commonly used machine learning technique for text analysis, latent Dirichlet allocation (LDA).

LDA is a generative probabilistic model for collections of discrete data such as text corpora. Introduced by Blei, Ng, and Jordan (2003), it is a three-level hierarchical Bayesian model where each item of a collection is modeled as a finite mixture over an underlying set of topics. Each topic is, in turn, modeled as an infinite mixture over an underlying set of topic probabilities.

LDA posits that documents (in this case, a patient entry for *reason for admission*) are composed of multiple topics, and a topic is a distribution over words in the lexicon. The fundamental assumption of LDA is that documents are mixtures of topics and that the topics are mixtures of words. We use LDA for its ability to uncover the underlying thematic structure of a document collection.

Before applying LDA, we first preprocess the data to make inference easier. Free-text notes in EMRs such as the *reason for admission* have many features that make them resistant to standard text analysis methods. For example, they are filled with medical jargon and medical abbreviations, do not follow standard sentence structure, often deviate from grammatical syntax, and have a very high rate of typographical errors. Keeping this in mind, we set our preprocessing parameters to be quite lax, in essence favoring a higher type II over type I error rate. In other words, we preprocess the data such that there is a high likelihood of retaining actual medical terminology with real information but also allows more "filler" words to influence analysis that would be the case if we had a stricter preprocessing protocol. To this end, we remove all punctuation, nonalphabetic characters, and stopwords (*and*, *or*, *if*, etc.) from the text field. We also lemmatize each word, i.e., converte it to its root form (e.g., *painful* and *paining* become *pain*). When creating our bag of words for the LDA analysis, we drop any words that occur in more than 95% of admissions (common words such as *yo*, signifying *year old* are unlikely to be meaningful) and any words that occur only once in the entire dataset (with the goal of excluding typos). Other than that, we make no restrictions.

We use LDA analysis to identify five primary themes. LDA does not identify what these themes are, but the five most common words from each theme are shown in the top panel in Figure A.8: For example, Topic 3 is associated with the words *right*, *left*, *back*, *hip*, and *surgery*. The visual distribution of the topics is also shown in the heatmap in the bottom panel. For example, Topic 4 is equally represented on average among Black and White patients, but Topics 1 and 2 are represented more among White patients and Topics 3 and 5 among Black patients. Regardless, the formal analysis in Table A.5 (which reestimates Equation 1a using each of the five topics as an outcome) finds that the distribution of topics does not change differentially with strain for Black vs. White patients, which once again serves as confirmation that differential patient selection with strain is not a concern or a threat to our identification assumption.

B Analysis of sentiment and detail

In this analysis, we use the TextBlob library, a Python-based natural language processing (NLP) tool, to perform sentiment analysis on a corpus of text data. TextBlob simplifies text processing in Python, providing an accessible interface for a range of NLP tasks including part-of-speech tagging, noun phrase extraction, and sentiment analysis. Of particular interest to this study is its application to evaluating the sentiment polarity and subjectivity of text data.

TextBlob's sentiment analysis function is grounded in a lexicon of words, where each word is associated with sentiment scores. When analyzing a given text, TextBlob calculates the overall sentiment by aggregating the sentiment scores of the words contained within the text. This process involves two primary components:

Subjectivity analysis: This quantifies the degree of personal opinion and factual information contained in the text. The subjectivity score is a reflection of the presence of personal views and subjective evaluations as opposed to factual, objective information, ranging from 0 to 1, where 0 is entirely objective and 1 is entirely subjective. Examples of hypothetical physician notes to illustrate variation in subjectivity scores are the following:

- "Patient expresses pain as a 3 out of 10" = Subjectivity score of 0
- "The patient's symptoms suggest possible risk of stroke" = Subjectivity score of 0.6

- "I am deeply optimistic that this patient may survive" or "Patient describing unimaginable levels of pain" = Subjectivity score of 0.9.

Subjectivity is assessed independently from polarity (described below, which identifies emotional tone), though there can be a correlation between the two scores. The third example above attempts to highlight the difference: A physician can write a subjective note that either conveys positive information or negative information. In our sample, the correlation between subjectivity and polarity is 0.15.

Polarity analysis: This is a measure of the emotional leaning of the text, indicating whether the expressed opinion in the text is positive, negative, or neutral. The polarity score provided by TextBlob is used to determine the sentiment orientation of each text entry in the dataset, ranging from -1 to 1, representing negative to positive sentiment. Examples of hypothetical physician notes to illustrate variation in polarity scores are as follows:

- "Patient responding exceptionally well to new treatment" = Polarity score of +0.7
- "Concerning lack of progress in patient's recovery" = Polarity score of -0.5
- "Patient received 5mg of medication" = Polarity score of 0.

It is important to note that like most basic sentiment analysis tools, TextBlob primarily analyzes the sentiment of individual words rather than sentiment based on syntactical structure or the context beyond immediate word combinations. For example, a physician note that stated "Concerned about lack of patient progress" would be assigned a similar polarity score to a note that stated "Patient is concerned about lack of progress," even though the notes have two different subjects and relay different information.

Adjectives: Finally, to measure the level of detail provided in the text data, we perform a rudimentary analysis where we simply count the number of adjectives in a given provider note. However, using ML techniques (such as the NLP toolkit) on physician notes is difficult because these algorithms are trained to identify adjectives based on the context in which they appear – such as the words before and after. As already discussed, physician notes largely do not follow syntax and grammar rules and so such ML techniques may not be reliable. Thus, instead, we rely on a very basic rule: We identify adjectives as words that end in [-*able*, -*al*, -*ant*, -*ary*, -*ful*, -*ic*, -*ish*, -*ive*, -*less*, -*ous*, -*y*, -*er*].

C Assessment of medical need by providers under strain

We examine strain-related racial disparities in mortality by triage score, motivated by the finding that Black patients in the ED are consistently perceived to be less sick, as indicated by lower triage scores, than White patients (Figure III). This lower perception of need contrasts with Black patients' higher mortality rates, which suggests there may be a "mismatch" between perceived and actual medical need.

We test the hypothesis that the increased mortality of Black relative to White patients at high strain stems from increasingly inaccurate clinical assessments. For the subsample of inpatients admitted through the ED, we reestimate Equation 1b, first with triage scores as the outcome (Figure A.6 Panel (a) and Table A.7 Column 1) and then with mortality (Figure A.6 Panel (b) and Table A.7 Column 2), adding a triple interaction between race, high strain, and patient medical need (using the Elixhauser mortality index median split from the entire inpatient sample).

Figure A.6 Panel (a) shows that the average triage scores are higher for high- than for low-need patients, both Black and White, and lower for Black patients than for White patients. Moreover, while high strain widens the racial mortality gap for high-need patients (Figure A.6 Panel (b)), no corresponding pattern in triage scores explains this pattern: For example, strain does not reduce the triage scores for high-need Black patients relative to those of high-need White patients. Consistent with this, the triple interaction in Table A.7 is not statistically significant. Therefore, we are unable to detect systematic strain-induced differences in provider assessments of medical need by race.

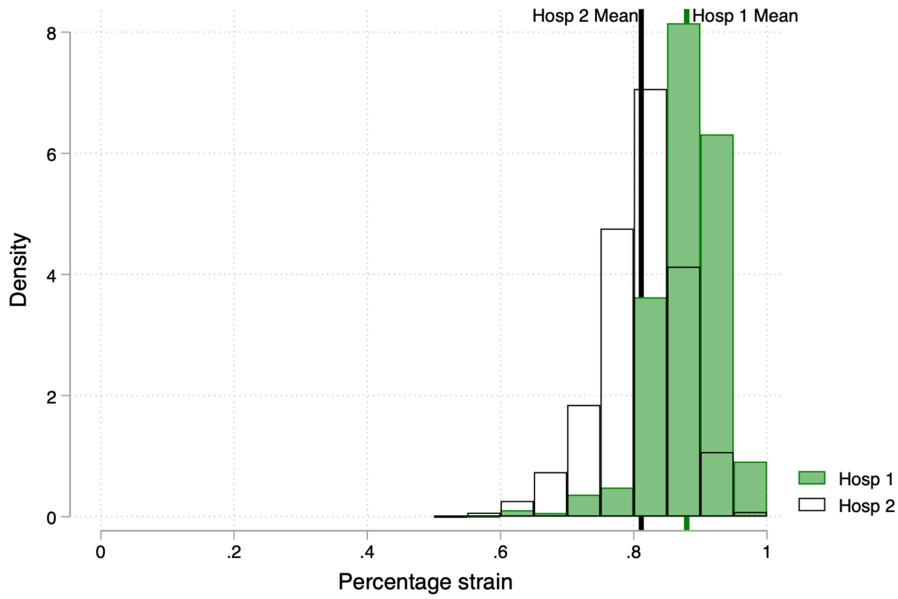
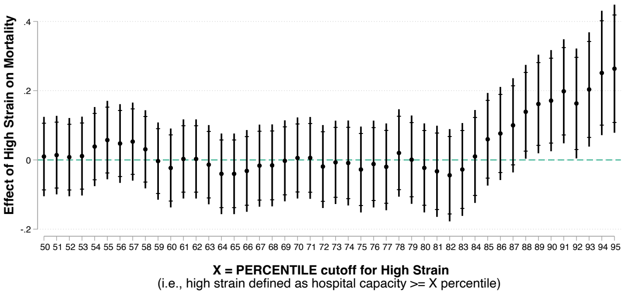
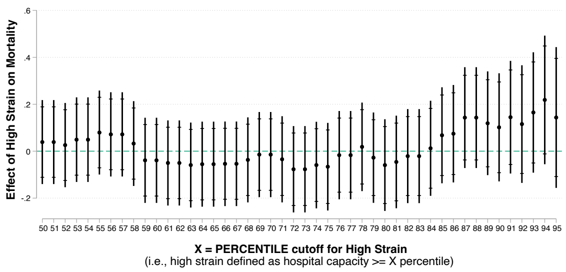


FIGURE A.1
Distribution of Capacity Strain by Hospital

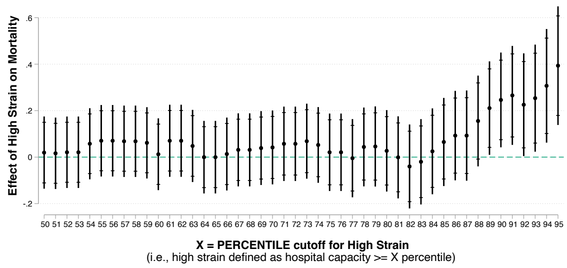
Note: This figure displays the distribution of hospital strain, defined as the percentage of inpatient beds occupied, separately for Hospital 1 (green) and Hospital 2 (white). Vertical lines indicate the mean strain levels for each hospital.



Panel (a): Full Sample



Panel (b): Hospital 1 (N= 43,834)



Panel (c): Hospital 2 (N= 63,390)

FIGURE A.2

Relationship Between Strain and Mortality

Note: Each estimate in a panel plots the coefficient from regressing in-hospital mortality on the [High strain] term from a separate regression. High strain equals 1 if hospital capacity at the patient's arrival hour is at or above percentile X, and 0 otherwise, where $X \in (50, 51, 52, \ldots, 95)$. Panel (a) reports results for the pooled sample; Panels (b) and (c) repeat the analysis for Hospital 1 and Hospital 2, respectively. Covariates include year, month, day-of-week, and hour-of-day, and major holiday fixed effect. Heteroskedasticity-robust 90% and 95% confidence intervals are shown.

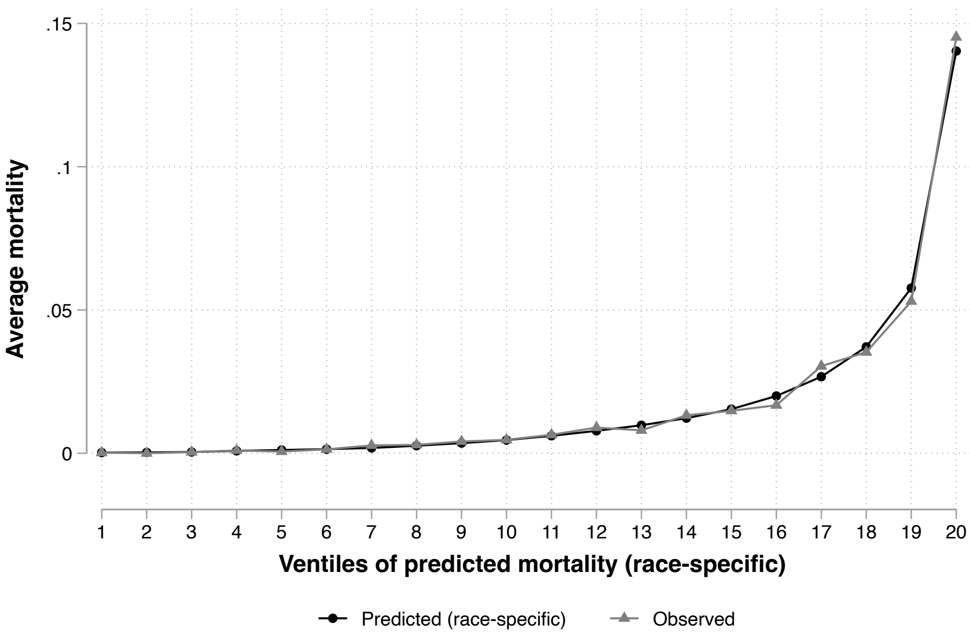
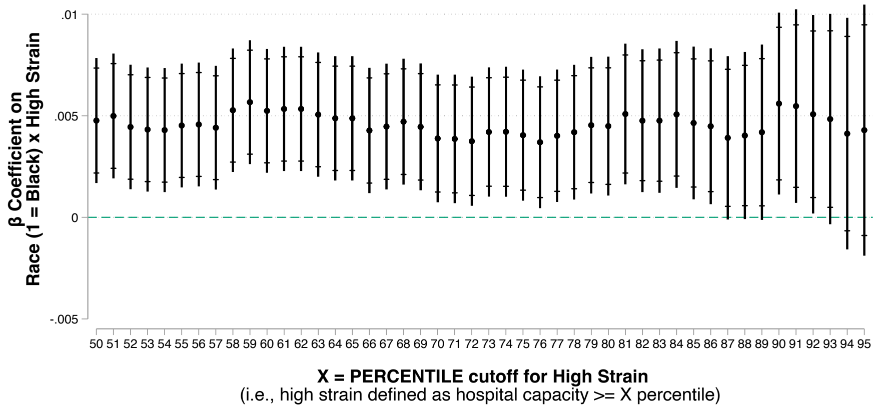


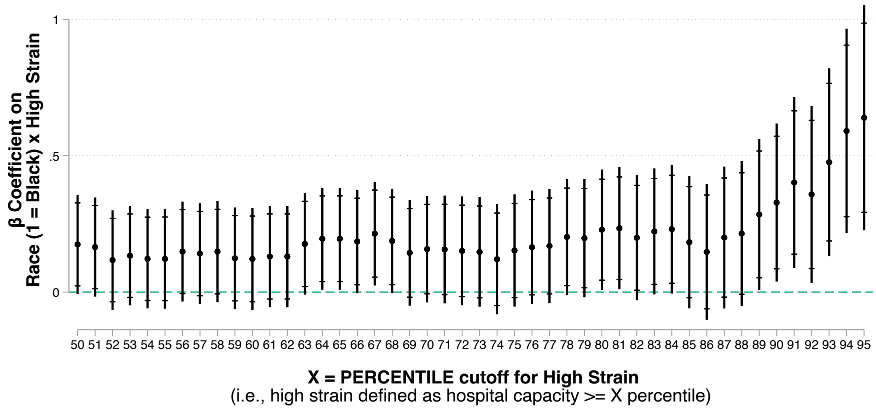
FIGURE A.3

Predictive Ability of (Race-Specific) Predicted Mortality Measure for Observed Mortality

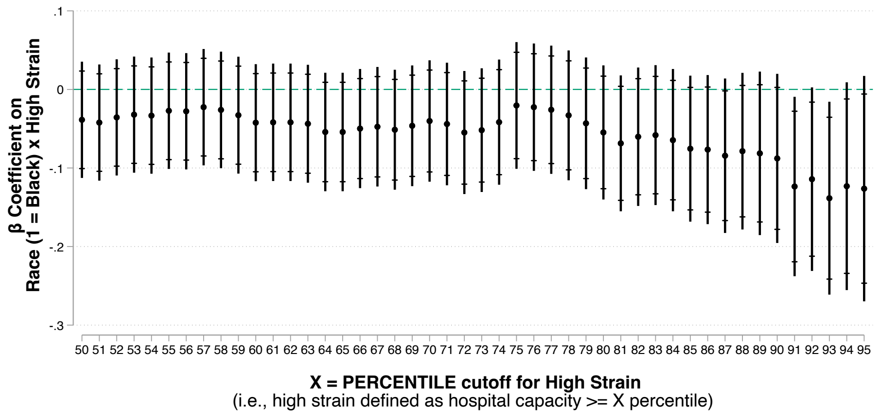
Note: Figure plots the average (race-specific) predicted mortality and the average observed mortality at each ventile of predicted mortality. The predicted (race-specific) mortality measure is estimated by logistic regressions of observed mortality on patient age, sex, insurance status, number of comorbidities, and Elixhauser mortality and readmission indices separately for Black and White patients.



Panel (a): In-hospital mortality



Panel (b): Wait times

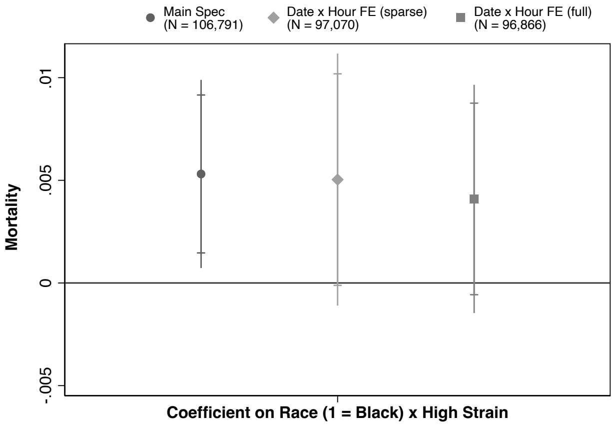


Panel (c): Provider effort index

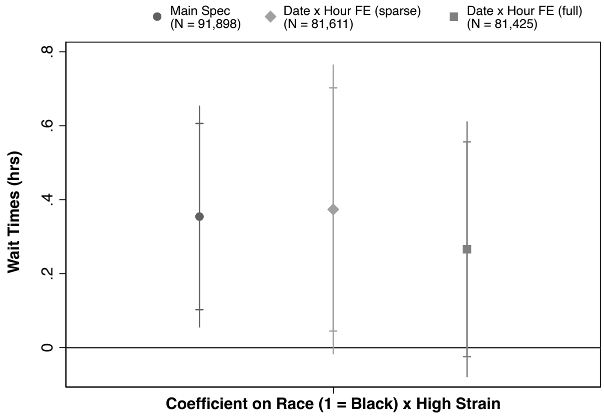
FIGURE A.4

Robustness to Definitions of High Strain at Alternative Percentile Cutoffs

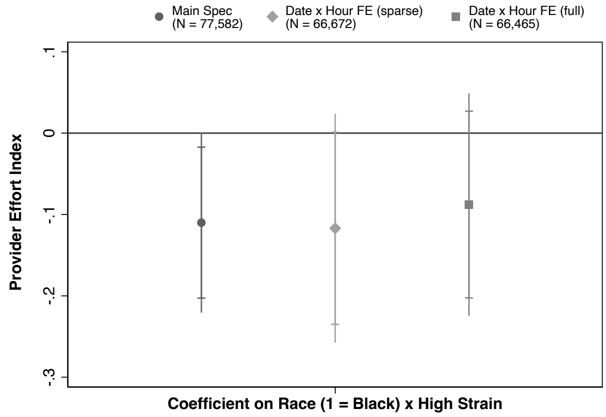
Note: Outcomes are: in-hospital mortality in Panel (a), wait times in Panel (b), and provider effort index in Panel (c). Each estimate in a panel plots the interaction coefficient on [Race (Black = 1) × High strain] from a separate regression of Equation 1b . High Strain equals 1 if hospital capacity at the patient's arrival hour is at or above percentile X, and 0 otherwise, where $X \in (50, 51, 52, \ldots, 95)$. Heteroskedasticity-robust 90% and 95% confidence intervals are shown.



Panel (a): Mortality



Panel (b): Wait times



Panel (c): Provider effort index

FIGURE A.5

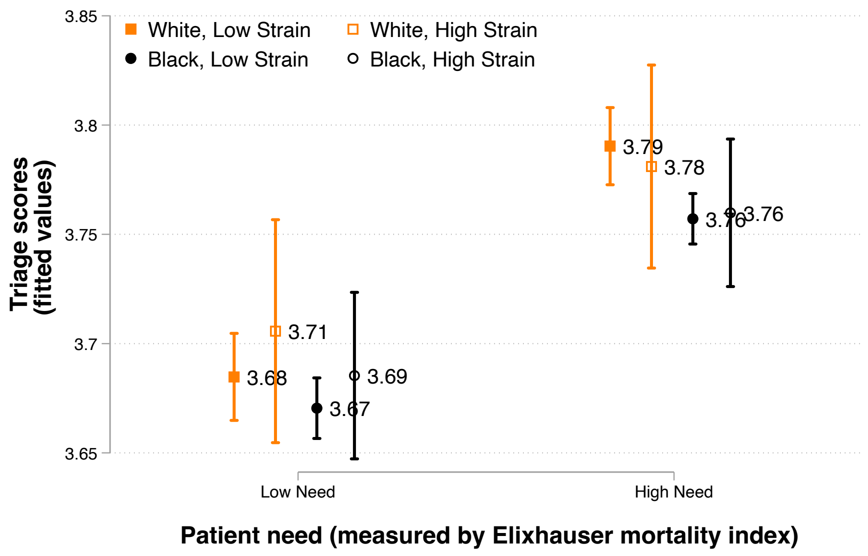
Robustness: Hospital × Date × Hour Fixed Effects

Note: Outcomes are: in-hospital mortality in Panel (a), wait times in Panel (b), and provider effort index in Panel (c). Heteroskedasticity-robust 90% and 95% confidence intervals are shown. Each estimate in a panel plots the interaction coefficient on [Race (Black = 1) × High strain] from one of the following 3 models:

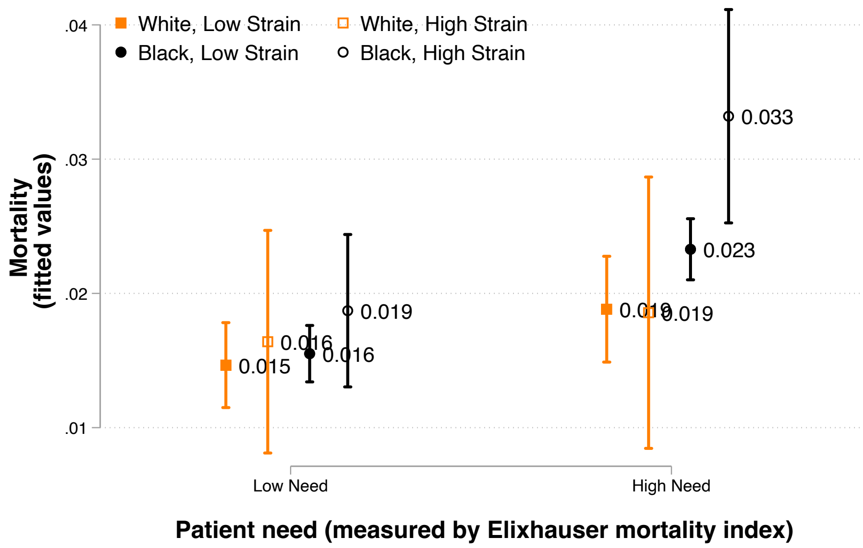
(1) *Main Spec:* Baseline model (Eq 1b)

(2) *Date × Hour FE (sparse)*: Baseline model (Eq 1b) with only race, high strain, and race x high strain on the RHS, as well as [hospital x year x month x day of week x hour of day] FEs replaced with 20,298 to 26,629 [hospital x arrival date x hour of day] FEs

(3) *Date × Hour FE (full)*: Baseline model (Eq 1b) with *all* variables on the RHS, as well as [hospital x year x month x day of week x hour of day] FEs replaced with 20,298 to 26,629 [hospital x arrival date x hour of day] FEs



Panel (a): Fitted values of triage scores
1 (lowest severity) - 5 (highest severity)



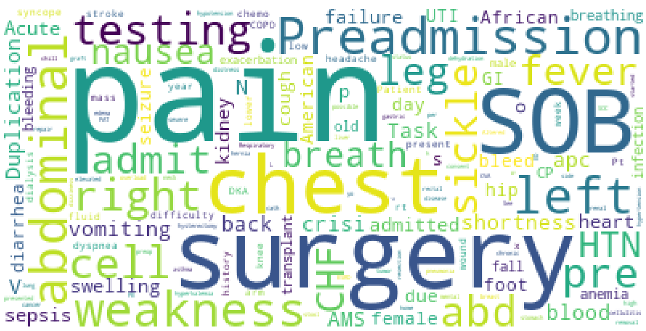
Panel (b): Fitted values of in-hospital mortality

FIGURE A.6

Test of Mismatch Between Patient Medical Need and Need as Assessed by Providers

Note: Outcomes are triage scores (discrete values of 1–5, reverse coded so higher values imply greater severity) in Panel (a) and in-hospital mortality in Panel (b) from estimating Equation 1b with all interactions between race (=1 for Black) × high strain (defined by the tenth decile) × high medical need (defined as above the inpatient sample median for Elixhauser mortality index). Estimated only on the sample of inpatients admitted through the ED. Estimates evaluated at the means of other covariates, with 95% confidence intervals based on robust standard errors. Regression coefficients are presented in Table A.7.

Panel (a): White patients

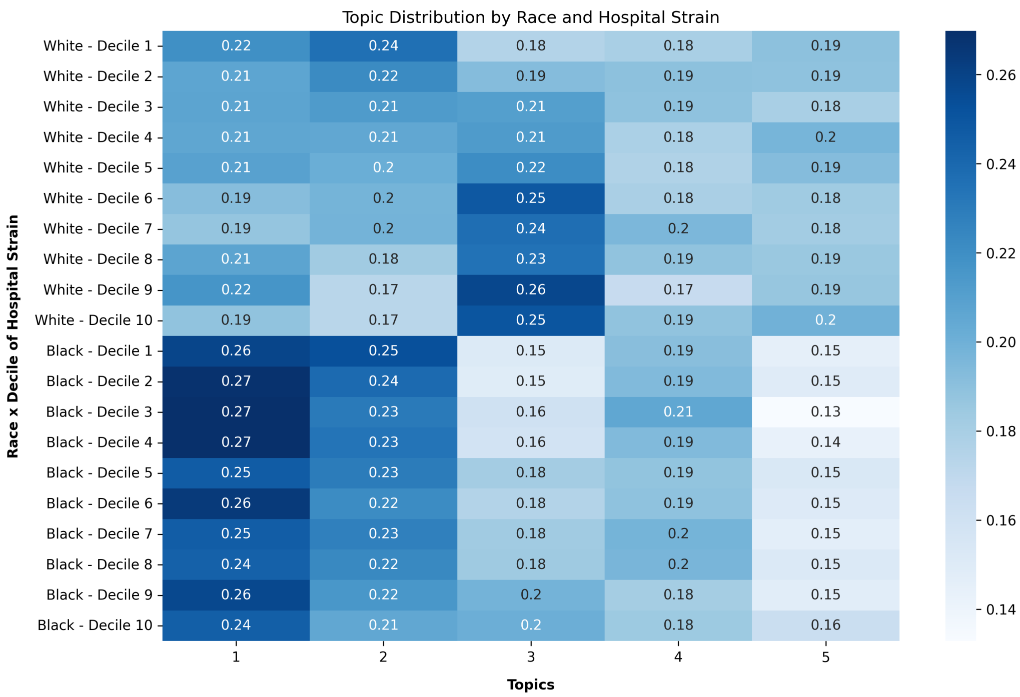


Panel (b): Black patients

FIGURE A.7
Word Clouds of *Reason for Admission* Note by Race

	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
Word 1	cell	duplication	right	am	breath
Word 2	preadmission	task	left	patient	right
Word 3	weakness	abdominal	back	yo	left
Word 4	testing	chest	hip	admitted	knee
Word 5	sob	pain	surgery	htn	fever

Panel (a): Themes identified by LDA

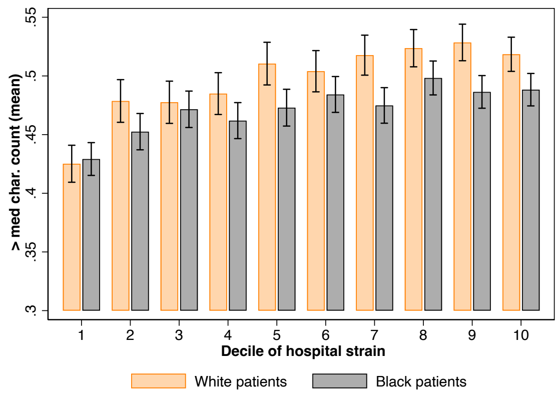


Panel (b): Heatmap of topic distribution

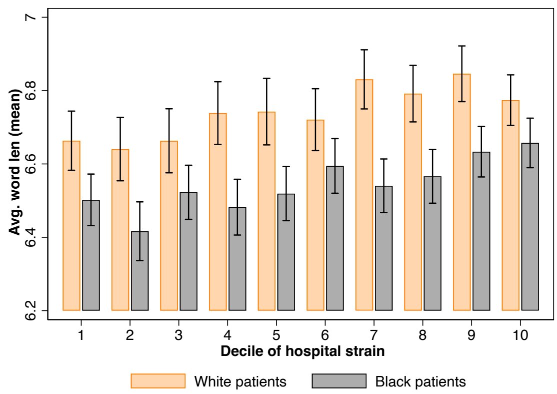
FIGURE A.8

Topic Identification and Distribution from Latent Dirichlet Allocation Analysis

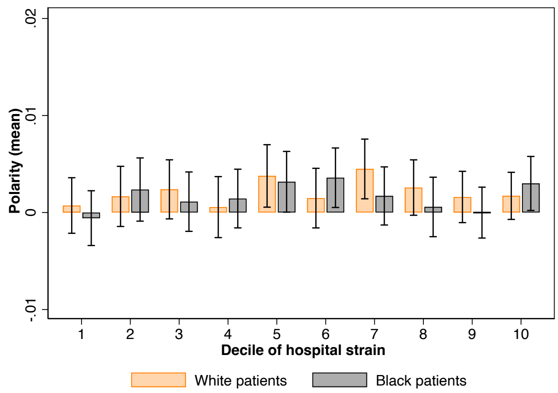
Note: Details about LDA are provided in Appendix A. Panel (a) identifies the five themes identified by LDA in the *reason for admission* note and the five most common words associated with each theme. Panel (b) provides a heatmap of the topic distribution by race and decile of hospital strain.



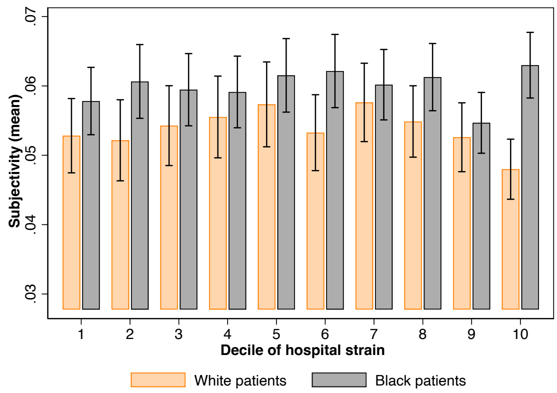
Panel (a): Character count



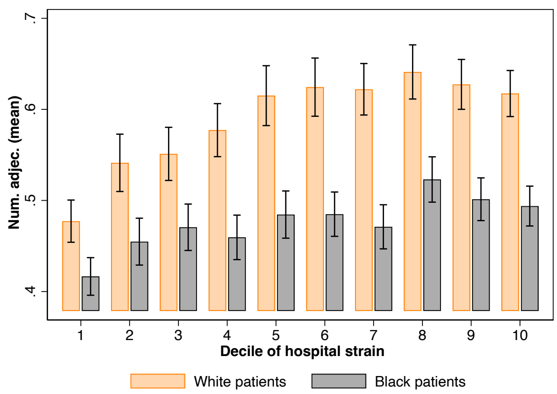
Panel (b): Word length



Panel (c): Polarity



Panel (d): Subjectivity

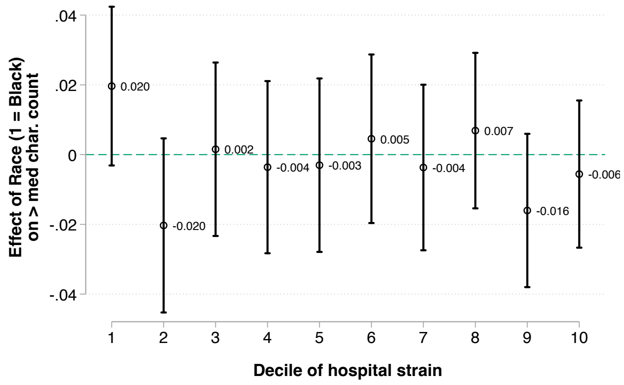


Panel (e): Adjectives

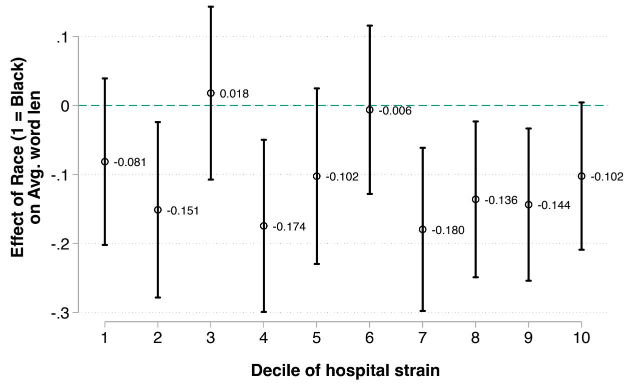
FIGURE A.9

Five (Text-Based) Measures of Provider Effort

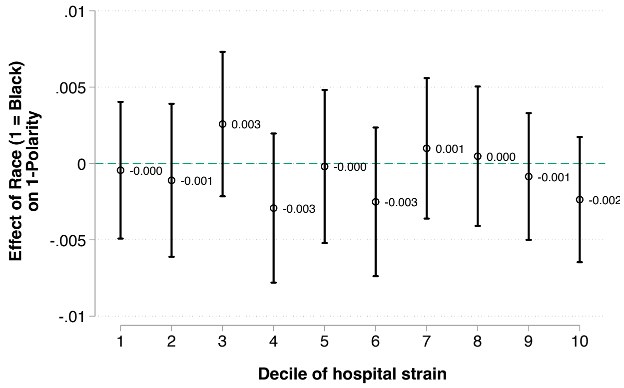
Note: Figure plots the raw means of the five different text-based features of the reason for admission note used to generate the summed provider effort index, by patient race and decile of strain. These include the (a) proportion above the sample median of character count, (b) average word length (in characters), (c) polarity (measured on a scale ranging from -1 to 1), (d) subjectivity (measured on a scale of 0 to 1), and (e) number of adjectives.



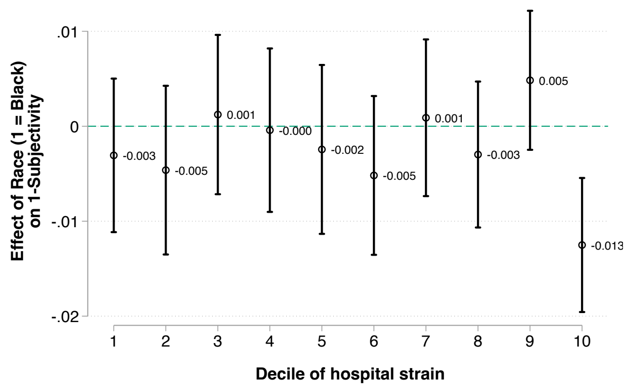
Panel (a): Marginal effect of race on character count



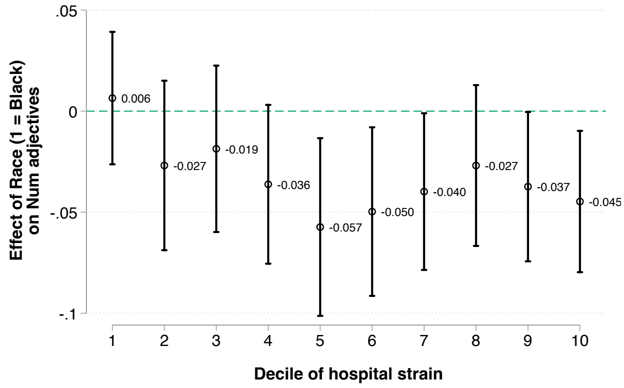
Panel (b): Marginal effect of race on word length



Panel (c): Marginal effect of race on polarity (reverse coded)



Panel (d): Marginal effect of race on subjectivity (reverse coded)



Panel (e): Marginal effect of race on number of adjectives

FIGURE A.10

Analysis of Each Component of the Provider Effort Index

Estimates from Equation 1a with each of the five effort measures derived from the *reason for admission* note (used to construct the summed provider effort index) as the outcome. Outcomes are: above-median character count (Panel (a)), average word length (Panel (b)), 1 - polarity (Panel (c); 1 - subjectivity (Panel (d)); number of adjectives (Panel (e)). All effects are evaluated at the means of other covariates. Heteroskedasticity-robust 95% confidence intervals are shown.

	Decile 1		Decile 2		Decile 3		Decile 4		Decile 5		Decile 6		Decile 7		Decile 8		Decile 9		Decile 10	
	H1	H2	H1	H2	H1	H2	H1	H2	H1	H2	H1	H2	H1	H2	H1	H2	H1	H2	H1	H2
Proportion full	0.69	0.78	0.75	0.84	0.78	0.85	0.80	0.87	0.81	0.88	0.82	0.89	0.84	0.90	0.85	0.91	0.87	0.93	0.91	0.95
	(0.04)	(0.05)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.02)	(0.01)
Observations	5758	6006	5803	3836	5641	3895	5902	3989	5768	3672	6462	3542	6508	3766	6556	4747	7464	4506	7281	5424

TABLE A.1

Hospital Capacity at Each Decile of Strain

Note: For hospitals H1 and H2, this table presents the mean proportion of beds filled at each of the ten deciles of strain, with SD rounded to 2 decimals points in parentheses.

	Dep Var: In-Hospital Mortality		
	(1) Pooled	(2) Hospital 1	(3) Hospital 2
Black	0.0019*	-0.0001	0.0043***
	(0.0010)	(0.0014)	(0.0013)
High strain (=dec 10)	-0.0005	-0.0020	0.0019
	(0.0019)	(0.0026)	(0.0026)
Black × High strain	0.0053**	0.0072*	0.0026
	(0.0023)	(0.0038)	(0.0031)
N	106,791	43,534	63,170
R^2	0.292	0.298	0.303
Sample mean	0.017	0.020	0.014

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.001$

TABLE A.2

Mortality Results by Hospital

Note: Coefficients from regressions of in-hospital mortality on race (=1 if Black) and the top decile of hospital strain (Eq. 1b). Column 1 reports estimates from the pooled sample. Columns 2 and 3 restrict to admissions at Hospital 1 and Hospital 2, respectively. Heteroskedasticity-robust standard errors in parentheses. The p-value for equality of the interaction coefficients across Columns 2 and 3 is equal to 0.35, and obtained from a fully interacted triple-difference model (Race × Strain × Hospital), in which all covariates are also interacted with hospital

	Dep Var: In-Hospital Mortality		
	(1)	(2)	(3)
Black	0.0012	-0.7931***	0.0009
	0.0008	0.1796	0.0010
High strain (dec=10)	0.0004	-0.0565	-0.0010
	0.0016	0.3219	0.0019
High medical need	0.0084***	1.0000***	
	0.0016	0.1489	
Black × High strain (dec=10)	0.0013	0.5688	0.0054**
	0.0018	0.4704	0.0024
Black × High medical need	-0.0009	0.7001***	
	0.0018	0.1861	
High strain (dec=10) × High medical need	-0.0014	0.0857	
	0.0035	0.3440	
Black × High strain (dec=10) × High medical need	0.0081*	-0.2884	
	0.0048	0.4952	
Objective symptoms			-0.0066*
			0.0036
Black × Objective symptoms			-0.0028
			0.0040
High strain (dec=10) × Objective symptoms			0.0169
			0.0115
Black × High strain (dec=10) × Objective symptoms			-0.0135
			0.0131
N	106800	107177	106800
R^2	0.284		0.284
Spec	Elix (LPM)	Elix (Poisson)	Symptom subjectivity (LPM)

TABLE A.3

Heterogeneity in Mortality by Patient Need/Subjectivity of Symptoms

Note: Estimates from Equation 1b with all main and interaction terms for Race x High strain x $[Var_i]$. Var_i is patient need (as measured by Elixhauser mortality index) for Columns 1 and 2, and objective symptoms (as measured by mentions of "chest pain" or "shortness of breath" in the *reason for admission* note) in Column 3. $p < .001^{***}$ $p < .05^{**}$ $p < 0.1^{*}$.

	(1) Age (yrs.)	(2) Female	(3) Uninsured	(4) # comorb	(5) Mort index	(6) Readm index
Black	-3.34***	0.10***	0.01***	-0.06***	-0.88***	1.66***
	(0.12)	(0.00)	(0.00)	(0.01)	(0.05)	(0.06)
High strain (=dec 10)	0.19	-0.00	0.00	0.01	-0.01	-0.14
	(0.25)	(0.01)	(0.00)	(0.02)	(0.11)	(0.12)
Black × High strain (=dec 10)	-0.07	0.01	-0.00	0.00	-0.04	0.22
	(0.29)	(0.01)	(0.01)	(0.02)	(0.13)	(0.14)
N	106791	106791	106791	106792	106800	106800
R^2	0.506	0.253	0.213	0.856	0.779	0.900
Mean of outcome	54.78	0.59	0.13	4.09	12.58	24.50

TABLE A.4

Tests for Patient Selection Using Individual Patient Covariates as Outcomes

Note: Estimates of Equation 1b with the following outcomes assessed in the full sample of inpatients: patient age (Col 1), patient sex (Col 2), insurance status (Col 3), total number of Elixhauser comorbidities (Col 4), Elixhauser mortality index (Col 5), and Elixhauser readmission index (Col 6). Covariates are the same as in Eq. 1b (except for the exclusion of the dependent variable in each column). $p < .001^{***}$ $p < .05^{**}$ $p < 0.1^{*}$

	(1) Topic 1	(2) Topic 2	(3) Topic 3	(4) Topic 4	(5) Topic 5
Black	0.018***	0.008**	-0.006*	-0.004	-0.016***
	(0.004)	(0.004)	(0.003)	(0.004)	(0.003)
High strain (=dec 10)	-0.015*	-0.007	0.004	0.017**	0.001
	(0.008)	(0.007)	(0.008)	(0.007)	(0.007)
Black × High strain (=dec 10)	-0.002	0.001	0.000	-0.012	0.014
	(0.010)	(0.009)	(0.009)	(0.009)	(0.009)
N	76156	76156	76156	76156	76156
R^2	0.199	0.247	0.264	0.199	0.214
Mean of outcome	0.23	0.21	0.20	0.19	0.17

TABLE A.5

Topic Analysis

Note: Regression coefficients from Equation 1b with the five topics (identified by the latent Dirichlet allocation) as outcomes. Details on LDA provided in Appendix A. $p < .001^{***}$ $p < .05^{**}$ $p < 0.1^{*}$

	(1) ICU Admission	(2) Length of stay	(3) Charges ($)
Black	-0.01***	0.03	-3416.52***
	(0.00)	(0.07)	(566.26)
High strain (=dec 10)	-0.02**	-0.05	-127.50
	(0.01)	(0.14)	(1095.53)
Black × High strain (=dec 10)	0.02**	0.22	1798.61
	(0.01)	(0.18)	(1295.75)
N	106791	106791	106768
R^2	0.411	0.276	0.423
Mean of outcome	0.24	6.57	52047.56

TABLE A.6
Wait Times and Other Inputs of Care

Note: Regression coefficients from Equation 1b with the following outcomes: ICU admission (Col 1), total length of inpatient stay in days (Col 2), and total inpatient charges (Col 3). $p < .001^{***}$ $p < .05^{**}$ $p < 0.1^{*}$

	(1)	(2)
	Triage score	In-hospital mortality
Black	-0.014	0.001
	(0.011)	(0.002)
High medical need	0.106***	0.004
	(0.014)	(0.003)
High strain (=dec 10)	0.021	0.002
	(0.027)	(0.004)
Black × High strain (=dec 10)	-0.006	0.001
	(0.032)	(0.005)
Black × High medical need	-0.019	0.004
	(0.015)	(0.003)
High strain (=dec 10) × High medical need	-0.030	-0.002
	(0.035)	(0.007)
Black × High strain (=dec 10) × High medical need	0.018	0.009
	(0.043)	(0.008)
N	49433	49433
R^2	0.272	0.361
Mean of outcome	3.728	0.020

TABLE A.7

Triage Score by Race, Need, and Strain

Note: Regression coefficients from estimating Equation 1b with all interactions between race (=1 for Black) × high strain (defined by the tenth decile) × high medical need (defined as above the sample median for Elixhauser mortality index). Outcomes are (i) Col 1: triage scores (1–5, reverse coded so higher values imply greater severity) and (ii) Col 2: in-hospital mortality. Estimated on the sample of inpatients admitted through the ED. $p < .001^{***}$ $p < .05^{**}$ $p < 0.1^{*}$.

	(1) Wait time (hrs)	(2) Provider effort
Black	0.98***	-0.14***
	(0.08)	(0.03)
High medical need	-0.41***	0.05
	(0.09)	(0.04)
High strain (dec=10)	0.65***	0.09
	(0.16)	(0.06)
Black × High strain (dec=10)	0.37*	-0.08
	(0.22)	(0.08)
Black × High medical need	-0.00	0.07*
	(0.10)	(0.04)
High strain (dec=10) × High medical need	0.82***	-0.06
	(0.20)	(0.08)
Black × High strain (dec=10) × High medical need	-0.10	-0.04
	(0.30)	(0.11)
N	91907	77593
R^2	0.292	0.218
Mean of outcome	5.46	0.00

TABLE A.8

Heterogeneity in Wait Times and Provider Effort by Patient Need

Note: Estimates from Equation 1b with wait times (Column 1) and provider effort index (Column 2) as outcomes and all main and interaction terms for race × high strain × patient medical need. High and low patient need measured by above- and below-median sample values for the Elixhauser mortality index. $p < .001^{***}$ $p < .05^{**}$ $p < 0.1^{*}$.

	Dep Var: In-Hospital Mortality	
	(1)	(2)
Black	0.0018*	0.0011
	(0.0011)	(0.0011)
High strain (dec=10)	-0.0008	-0.0007
	(0.0020)	(0.0021)
Black × High strain (dec=10)	0.0065**	0.0070**
	(0.0026)	(0.0028)
N	91898	77582
R^2	0.306	0.315
Mean of outcome	0.0186	0.0172

TABLE A.9

Mortality: Wait Time and Provider Effort Subsamples

Note: Regression coefficients from Equation 1b for in-hospital mortality using only those observations for which data on wait time (Col 1) or provider effort (Col 2) are available. $p < .001^{***}$ $p < .05^{**}$ $p < 0.1^{*}$.

	(1) Provider effort	(2) > med chars	(3) Avg. word len	(4) 1-Polarity	(5) 1-Subjectivity	(6) Num adjectives
Black	-0.096***	-0.001	-0.109***	-0.000	-0.001	-0.030***
	(0.022)	(0.004)	(0.023)	(0.001)	(0.002)	(0.007)
High strain (=dec 10)	0.064	0.005	0.008	-0.000	0.008**	0.006
	(0.045)	(0.009)	(0.046)	(0.002)	(0.003)	(0.016)
Black × High strain (=dec 10)	-0.110*	-0.004	0.009	-0.002	-0.011**	-0.014
	(0.056)	(0.011)	(0.057)	(0.002)	(0.004)	(0.019)
N	77582	77582	77582	77582	77582	77582
R^2	0.220	0.227	0.172	0.142	0.153	0.241
Mean of outcome	0.003	0.483	6.637	0.998	0.943	0.528

TABLE A.10

Provider Effort Index (Summed, and Components)

Note: Regression coefficients from Equation 1b. Col 1 uses the main provider effort index as the outcome. The outcomes for the remaining columns are the five individual components of this index, as follows: above the median character count (Col 2), average word length (Col 3), 1 - polarity (Col 4), 1 - subjectivity (Col 5), and number of adjectives (Col 6). Descriptions for each of these outcomes are provided in Section 5. $p < .001^{***}$ $p < .05^{**}$ $p < 0.1^{*}$.