Deciding Under Pressure: Time Constraints and Professional Decision-Making

Ashley Litwin*

February 9, 2026

Abstract

Many high-stakes jobs require workers to acquire and process information under time constraints. When time is scarce, how do professionals implement formal decision rules, and which information do they rely on when time pressure makes those rules costly to apply? I study these questions in emergency department (ED) triage, where nurses implement a standardized protocol to prioritize patients for care. Exploiting quasi-random, minute-to-minute fluctuations in available decision time, I estimate within-nurse, within-shift responses using nearly two million triage decisions by 1,800 nurses across six hospitals. Under time pressure, nurses acquire less case-specific information: they spend less time evaluating patients, collect fewer diagnostic signals, and produce less detailed documentation. These changes systematically alter how nurses apply the protocol, placing less weight on case-specific information and more on coarse, readily observable patient characteristics, including age, race, and gender. While age enters the protocol as a coarse indicator of illness, race does not; yet reliance on race rises when time pressure makes the protocol hardest to apply, reflecting substitution toward observable cues shaped by recent experience. These deviations propagate downstream, generating group differences in wait times, diagnostic assessment, and treatment intensity, including for time-sensitive conditions such as stroke. Together, the results show how organizational constraints can weaken the disciplining role of standardized protocols and reshape high-stakes decisions, creating systematic disparities in outcomes even when formal rules and incentives remain unchanged.

Keywords: Time pressure; information; decision rules; organizations; labor; healthcare.

*Department of Business Economics and Public Policy, The Wharton School, University of Pennsylvania, Philadelphia, PA 19104; alitw@wharton.upenn.edu. For their support and guidance, I thank Judd Kessler, Corinne Low, and Alex Rees-Jones. For helpful comments and suggestions, I thank Diane Alexander, Alex Chan, Zoe Cullen, Christine Exley, Marius Guenzel, Atul Gupta, Alex Imas, Sasha Indarte, Ryan Oprea, Heather Schofield, and Kevin Volpp. Stephanie Dodd provided excellent research assistance. Funding came from the Wharton Behavioral Lab, the Wharton Coalition for Equity and Opportunity, and the Wharton Environmental, Social, and Governance Initiative. This project received approval from the University of Pennsylvania Institutional Review Board in September 2023.

1 Introduction

Many organizations rely on formal protocols to guide high-stakes decision-making, in part to promote consistency when time and attention are scarce. Yet the very time pressure that motivates such protocols can also constrain how they are applied in practice. Across a range of settings—police officers making use-of-force decisions, institutional investors evaluating trades, and doctors diagnosing and treating patients—high-skilled workers must act quickly and often with incomplete information, limiting their ability to fully acquire, process, and integrate the information prescribed by formal procedures. In principle, time pressure could lead decision-makers to apply formal rules more mechanically, or instead to substitute toward simpler, immediately observable inputs. Despite the prevalence of protocol-driven, time-sensitive decisions, we know little about which of these responses dominates in practice.

In this paper, I study how acute time pressure reshapes the application of formal decision rules in a high-stakes organizational setting. Identification requires observing the same decision-maker evaluating comparable cases under different time constraints, while holding constant the underlying characteristics of the decision problem—variation that is rarely available in practice. I overcome this challenge by exploiting minute-to-minute fluctuations in time pressure that are plausibly orthogonal to case characteristics, allowing clean estimation of how acute time constraints alter information acquisition and how available information is translated into decisions.

I examine these dynamics in the context of emergency department (ED) triage, where nurses must rapidly evaluate incoming patients and assign treatment priority using standardized protocols. Using detailed, time-stamped administrative data on 1.9 million triage decisions by 1,800 nurses across six hospitals between 2017 and 2025, I exploit short-run variation in time pressure measured at the onset of each triage encounter. These fluctuations arise from stochastic patient arrivals and variation in the duration of immediately preceding assessments, generating differences in available evaluation time that are not mechanically related to patient characteristics or clinical urgency. This setting provides high-frequency variation in available evaluation time that allows me to study how time pressure affects information acquisition and decision rules within nurse–shift comparisons.

I begin by showing that acute time pressure compresses information acquisition at the point of decision. As time pressure increases, nurses spend less time evaluating patients, collect fewer diagnostic signals, and produce shorter clinical notes. Across three quasi-random measures of time pressure—the number of patients waiting to be triaged, the current wait time of the next patient in the triage queue, and the duration of the immediately preceding triage encounter—higher pressure leads to systematic reductions in the amount of case-specific information incorporated into the decision. These results establish that short-run time constraints directly alter the information environment in which professional decisions are made.

When individualized information is compressed, nurses systematically reweight triage decisions toward coarse, readily observable patient characteristics. I show that as time pressure increases, the estimated weight placed on patient age, gender, and race—features that are always observed at the moment of decision—rises systematically in triage decisions. Differences in triage priority across these characteristics widen monotonically as constraints tighten. These effects are identified within nurse–shift and hospital–hour and hold fixed a rich set of clinical covariates, and therefore reflect changes in decision weights rather than changes in patient composition. The magnitude of these effects is substantial: moving from low- to high-pressure states increases the implied weight on race and age by roughly 3–4 percentage points, or 50–60 percent relative to baseline, with smaller absolute increases for gender. Importantly, these results go beyond documenting information compression under time constraints: they show that acute constraints systematically reweight which inputs enter high-stakes decisions.

That race is reweighted alongside age under time pressure raises a natural puzzle. Increased reliance on age under information compression is consistent with triage protocols, which explicitly incorporate age as a coarse indicator of physiological risk. By contrast, triage guidelines are race-neutral: race plays no role in the mapping from observed symptoms and vital signs to triage priority. Why, then, does race become more influential in triage decisions precisely when individualized information is most limited?

To address this puzzle, I next examine how race enters simplified, informal decision rules in practice. A plausible mechanism is experience-based learning: repeated exposure to similar

cases can generate coarse associations that guide judgment under constraint (Mullainathan, Schwartzstein and Shleifer, 2008; Malmendier and Nagel, 2011, 2016). In the ED, frequent users—often perceived as low urgency—provide a salient source of such experience. Because frequent ED use is unevenly distributed across White and non-White patients, race becomes a dimension along which experience-based associations may form. Consistent with this, I show that nurses' recent exposure to frequent ED users systematically shapes how race enters subsequent triage decisions: greater exposure to non-White frequent users is associated with larger racial gaps in triage priority, while exposure to White frequent users attenuates or reverses these gaps. These patterns indicate that the simplified, informal decision rules used under time pressure reflect learning from accumulated experience rather than fixed underlying preferences or beliefs.

I then test how these experience-based shortcuts are applied once activated. Even if such rules are learned from relevant past cases, they may spill over to subsequent decisions with little case-by-case selectivity. Using short-run sequencing within the triage queue, I show that patients who happen to follow a frequent ED user—whose identity contains no information about the urgency of the next patient—face systematically different decision weights, with the direction of the effect depending on the race of the preceding frequent user. Because queue order is mechanically determined by arrival timing within a shift, these patterns are inconsistent with careful inference about the current patient's condition. Instead, they indicate that simplified, informal decision rules persist beyond the situations that motivate their use.

Finally, I examine how distortions in triage decisions propagate into downstream outcomes. Because triage priority structures subsequent clinical workflows, differences at this initial stage can affect how quickly patients are evaluated and treated. Consistent with this, visits evaluated under greater time pressure exhibit larger racial gaps in waiting times, diagnostic assessment, and treatment intensity, including for time-sensitive conditions such as stroke. These results suggest that small distortions in early decision rules can cascade through subsequent stages of care.

Taken together, the results show that even short-run constraints on decision time can meaningfully reshape individual decision-making inside protocolized organizations. Acute

time pressure compresses individualized information acquisition and induces substitution from formally prescribed decision rules toward simplified, experience-based rules that—once formed—spill over across subsequent cases. More broadly, the findings highlight how localized frictions faced by individual decision-makers can generate systematic and predictable deviations from formal protocols, producing distributional consequences even when rules, incentives, and personnel remain unchanged. The implications extend beyond emergency medicine to a wide range of settings in which organizations rely on formal procedures but require rapid individual judgment.

This paper makes three main contributions to the literature. First, it contributes to the organizational economics literature studying the gap between formal organizational design and realized decision-making. A central concern in this literature is why organizations often fail to achieve the behavior prescribed by formal procedures, even when incentives, monitoring, and personnel remain unchanged. Prior work often emphasizes agency problems, misaligned incentives, or limited enforcement as sources of deviation from formal rules (Bandiera, Barankay and Rasul, 2007; Olken, 2007; Blonz, 2023). This paper highlights a complementary mechanism: acute constraints on decision time, arising from organizational design, induce frontline workers to substitute away from formally prescribed decision rules toward simplified, experience-based rules. By exploiting minute-to-minute variation in time pressure within the same decision-maker, the paper provides direct evidence that localized frictions can systematically reshape decision rules in practice, generating predictable deviations from organizational protocols even in the absence of changes in incentives, oversight, or worker composition.[1]

Second, the paper contributes to the behavioral economics literature on coarsened decision-making by identifying acute time pressure as a key constraint that endogenously activates simplified decision rules. A growing body of theoretical and empirical research documents that decision-makers rely on coarse representations of information rather than fully individ-

[1]The paper also relates to the economics literature on clinical decision-making, which documents that physician behavior responds to fatigue over shifts (Jin et al., 2020), recent experience (Singh, 2021; Jin et al., 2023), peer behavior (Silver, 2021), and cognitive load or case complexity (Handel et al., 2025). Related work shows that time pressure can reduce diagnostic documentation and alter follow-up care (Freedman et al., 2021). The present paper complements this literature by isolating acute time pressure as a decision environment that truncates information acquisition itself and by examining how decisions are made once information use is constrained.

ualized assessment in settings including courts, labor markets, and finance (Mullainathan, Schwartzstein and Shleifer, 2008; Bordalo et al., 2025; Chen, Moskowitz and Shue, 2016; Radbruch and Schiprowski, 2024; Barberis, Shleifer and Vishny, 1998; Akepanidtaworn et al., 2023). This paper advances that literature by providing large-scale field evidence on the microfoundations of coarsened decision-making in a high-stakes professional setting. Exploiting minute-to-minute variation in decision time, it shows when simplified decision rules are activated by acute time constraints, how their content is shaped by accumulated experience, and that—once formed—they persist and spill over across decisions even when the cues that motivate their use are not relevant to the current case.[2]

Third, the paper contributes to the economics of discrimination by highlighting the role of decision environments in generating unequal treatment. Much of the existing literature focuses on distinguishing between belief-based and preference-based explanations for disparities across groups (Bertrand and Duflo, 2017; Lang and Kahn-Lang Spitzer, 2020; Litwin and Low, 2025).[3] While this distinction is central in many contexts, the results here emphasize a complementary mechanism: acute time pressure compresses information acquisition and shifts decision weights toward coarse, experience-based cues, generating disparities through changes in how decisions are made rather than through differences in beliefs and preferences alone. This mechanism yields testable predictions about the direction and severity of unequal treatment, with disparities emerging and intensifying along dimensions that become salient through experience and are embedded in simplified decision rules under pressure. In this sense, the results align with constructivist perspectives that view group characteristics not as fixed inputs into decision-making, but as categories whose relevance is inferred from experience and context (Rose, 2023). By identifying this mechanism at the point of decision-making, the paper provides a micro-level foundation for downstream disparities in care and outcomes documented in this and other high-stakes institutional settings (Singh and Venkataramani, 2024).[4]

[2]Laboratory studies show that time pressure affects risk-taking, strategic reasoning, cooperation, and decision quality (Kocher, Pahlke and Trautmann, 2013; Lawson, Larrick and Soll, 2020; Imas, Kuhn and Mironova, 2022; Kessler et al., 2025); this paper complements that work by identifying the organizational decision-rule channel through which time constraints reshape behavior in the field.

[3]A review by Bohren et al. (2023) finds that 62% of discrimination papers published in top economics journals from 1990–2018 discuss this distinction.

[4]Singh and Venkataramani (2024) document that as hospitals approach capacity, in-hospital mortal-

The remainder of the paper is organized as follows. Section 2 describes the institutional setting and data and Section 3 outlines the empirical strategy. Section 4 details the main results, while Section 5 discusses downstream implications of informal decision rules. Section 6 concludes.

2 Institutional Setting and Data

I study triage decisions for the universe of patients who arrived at one of six EDs within a large U.S. health system between January 2017 and April 2025.[5] The system's hospitals are non-profit and include urban teaching hospitals as well as other acute-care hospitals serving nearby suburban communities, all of which use a unified electronic health record and standardized triage protocol. The analysis draws on detailed administrative data extracted from the electronic health record, which captures every clinical and operational data point entered by staff in real time, providing a near-census of ED activity.

My dataset comprises approximately 1.9 million triage encounters conducted by roughly 1,800 unique nurses. Each visit record includes pre-triage patient characteristics (age, sex, race and ethnicity, arrival mode, and chief complaint), detailed information collected during triage (vital signs, triage notes, pain score, and the assigned priority level), and post-triage outcomes (waiting times, diagnostics, treatments, and discharge diagnosis). Crucially, the data contain precise timestamps for every key event—including patient arrivals, triage start and end times, and downstream encounters—which allow me to reconstruct the complete sequence of arrivals and triage decisions for each nurse and shift. This rich temporal structure links each decision to preceding and subsequent activity in the department and enables me to construct high-frequency measures of time pressure, information acquisition, and decision-making.

ity rises for Black but not White patients, consistent with race-based rationing of inpatient beds. While these patterns are consistent with multiple underlying channels, the present paper complements this evidence by identifying a micro-level mechanism at the point of decision-making: as acute time pressure increases, decision-makers substitute toward coarse, readily observable characteristics, reweighting race and other demographic cues in ways that can plausibly generate the downstream disparities in care and outcomes documented in that setting.

[5] Use of the de-identified administrative data was approved by the University of Pennsylvania Institutional Review Board.

Upon arrival to the ED, each patient first completes a brief registration process in which basic information — age, sex, race and ethnicity, arrival mode, and presenting complaint — is recorded by intake staff. Patients then wait in a queue to be evaluated by the triage nurse, who calls them in order of arrival as capacity allows. Patients are effectively assigned to whichever nurse becomes available next, so assignment is quasi-random conditional on arrival time and overall department busyness.[6] During the triage encounter, the nurse measures vital signs and completes standardized screening questions — for example, assessing fall risk, pain level, or recent travel—and records brief clinical notes that synthesize the nurse's assessment of the patient's condition and provide context to downstream providers. The nurse then assigns an Emergency Severity Index (ESI) level, a standardized five-point scale used nationally in U.S. EDs to classify patients by clinical urgency: 1 indicates life-threatening emergencies requiring immediate attention, and 5 denotes non-urgent visits that can safely wait.[7]

The triage decision is the first formal point of clinical judgment and directly affects how quickly patients are roomed, evaluated by a physician, and begin diagnostic testing and treatment. The assessment typically lasts only a few minutes and occurs continuously throughout a 12-hour shift as patients arrive. Because arrivals are inherently stochastic, the conditions under which each triage occurs varies naturally throughout a shift, creating short-run fluctuations in the time pressure nurses face. I leverage this variation in my empirical analysis outlined in Section 3.

2.1 Sample Construction and Key Variables

The unit of observation is a patient visit to the ED. Encounters with missing nurse or hospital identifiers, implausible timestamps, or incomplete information on patient age, race, or gender are excluded.[8] After applying these restrictions, the analysis sample includes approximately

[6] Although triage nurses have discretion to call visibly ill patients ahead of others, hospital protocol is to triage patients in order of arrival. In practice, approximately 15 percent of encounters in my data occur out of sequence. All results are robust to including fixed effects for whether a patient was skipped and to excluding out-of-sequence cases altogether.

[7] The Emergency Severity Index (ESI) is used by approximately 94 percent of U.S. hospitals (Worth et al., 2019).

[8] I trim implausible values of time-in-triage and waiting time at the 5^{th} and 95^{th} percentiles; results are qualitatively unchanged when using the full sample or alternative trimming rules.

1.9 million visits conducted by 1,800 triage nurses across six hospitals.[9]

Patient Demographics Patient race, gender, and age are defined using self-reported information in the electronic health record. Race and ethnicity are collapsed into a binary indicator for non-Hispanic White versus non-White.[10] Age is measured in years at the time of the emergency department visit and, in the main analysis, is coded as an indicator for age 65 and older, capturing a coarse distinction commonly used in clinical settings that may plausibly enter decision-making under time pressure.

Time Pressure I define time pressure as the momentary time constraint the nurse faces at the start of each triage assessment. Because triage encounters occur sequentially and patient arrivals vary unpredictably, this urgency fluctuates from minute to minute within a shift. I capture this variation using three real-time proxies measured at the start of each triage: (i) the number of patients waiting to be triaged behind patient i, (ii) the current waiting time of the *next* patient in line $(i+1)$ when patient i begins triage, and (iii) the duration of the immediately preceding triage encounter $(i-1)$. These measures provide complementary proxies for short-run time constraints and generate high-frequency within-shift variation in the environment in which each triage decision is made.

Information Acquisition I capture how much information the nurse collects and processes during triage along three complementary dimensions. First, I use the duration of the triage encounter, measured in minutes, which reflects the time spent gathering and processing clinical information. Second, I record the number of structured fields completed during triage — vital signs (temperature, pulse, respiratory rate, oxygen saturation, and blood pressure) and standardized screening questions — out of twelve items, reflecting the extent of diagnostic information acquired. Third, I assess information processing using the length of the free-text triage note, defined as the log number of words recorded, which reflects the level of narrative detail the nurse records during the assessment. Together, these measures proxy the degree of individualized assessment.

[9]Less than 5 percent of observations are excluded due to missing demographic information.

[10]This approach aligns with common practice in the health disparities literature. See, for example, Alsan, Ianni and Peterson (2024) and Singh and Venkataramani (2024).

Table 1: Summary Statistics

	(1) Mean	(2) SD	(3) N
Panel A. Patient Characteristics			
Age	47.47	21.78	1,896,860
Male	0.44	0.50	1,896,860
Non-Hispanic White	0.41	0.49	1,896,860
Arrived by ambulance	0.21	0.40	1,896,860
Panel B. Nurse-Shift Context			
Patients per nurse-shift	24.13	20.13	1,896,860
Panel C. Triage Assessment			
Top triage priority (ESI 1–2)	0.24	0.43	1,896,860
Minutes in triage	3.49	2.37	1,716,349
Clinical signals collected (out of 12)	8.89	3.40	1,896,860
Note length (words)	37.48	35.00	1,896,860
Panel D. Time Pressure			
Patients waiting	1.19	1.69	1,810,631
Next-patient wait time (min)	6.52	12.64	1,810,631
Duration of prior triage (min)	3.55	2.29	1,415,765

Note: The unit of observation is a patient visit to the emergency department. The table reports means, standard deviations, and non-missing counts. "Arrived by ambulance" equals one for patients transported by EMS (ambulance, fire rescue, police, and air/ground medical transport). "Patients per nurse–shift" is the total number of patients triaged by a nurse during a 12-hour shift. "Minutes in triage", "clinical signals collected", and "note length" summarize information recorded during the triage encounter. Time pressure proxies are measured at the start of triage; counts for these variables are lower when queue positions or adjacent-triage durations cannot be computed (e.g., the first patient of a shift or invalid timestamps).

Triage Decisions and Decision Weights Triage priority is assigned by the nurse at the end of the triage encounter and reflects a weighted assessment of multiple inputs, including observed patient characteristics and collected clinical information. In practice, this decision aggregates both individualized signals—such as vital signs, symptoms, and medical history—and coarse, readily observable characteristics. The empirical analysis described in Section 3 exploits time-pressure–induced variation in information acquisition to study how the relative weight placed on these inputs in the triage assignment itself shifts under constraint, holding fixed patient composition and clinical presentation. The main outcome variable is an indicator for receiving top triage priority (ESI level 1 or 2), which designates that a patient either requires immediate life-saving intervention (ESI 1) or is considered high-risk and warrants

rapid evaluation (ESI 2).

Table 1 presents summary statistics for the analysis sample. The average patient is 47 years old, 44 percent male, and 59 percent non-White, and 21 percent arrive by ambulance. Triage nurses evaluate roughly 24 patients per 12-hour shift, providing ample scope for within-shift identifying variation; in the busiest 10 percent of shifts, nurses evaluate more than 55 patients. During the triage assessment, nurses spend an average of 3.5 minutes with each patient, complete 8.9 of the twelve structured clinical fields, and record notes averaging 37 words, indicating that the triage encounter is both brief and information-dense. The time-pressure measures—patients waiting, the next patient's waiting time, and the duration of the prior triage—exhibit substantial dispersion across encounters, reflecting the highly dynamic decision environment in which these assessments occur.

3 Empirical Strategy

This section describes how I identify the effect of acute time pressure on information use and the mapping from patient inputs to triage decisions. The empirical strategy exploits minute-to-minute fluctuations in nurses' available decision time that arise from the stochastic arrival of patients to the ED. By comparing patients evaluated by the same nurse within the same shift under slightly different time constraints, the design isolates short-run variation in time pressure while holding fixed the broader clinical environment, institutional workload, and nurse-specific decision-making tendencies.

3.1 Identifying Variation in Time Pressure

I exploit three complementary, quasi-random sources of variation in the time available for each triage. The first is the number of patients waiting to be triaged when the nurse begins the current evaluation. A larger queue increases urgency and may induce nurses to move more quickly through each assessment. Conditional on hospital-by-hour and nurse–shift fixed effects, short-run fluctuations in queue length are driven by random patient arrivals and are plausibly orthogonal to the clinical severity of the current patient.

The second source of variation holds queue length constant and instead exploits the

waiting time of the next patient in line. Even when the number of patients waiting is the same, pressure to move faster is greater when the next patient has already been waiting longer. For example, two patients may both be triaged when two others are waiting behind them, but if the next person in line arrived ten minutes earlier rather than three minutes, the nurse faces greater urgency to complete the current assessment. Because the next patient's wait time depends on the stochastic timing of prior arrivals, it varies independently of the current patient's characteristics and provides a within-queue measure of time pressure.

The third source of variation captures the duration of the immediately preceding triage. When a nurse spends unusually long on the prior patient, she may compensate by abbreviating the subsequent evaluation, reducing the time available for information acquisition in the current encounter. This variation reflects idiosyncratic differences in how long earlier assessments happened to take—such as documentation or stabilization needs—rather than systematic differences in the current patient's condition.

Because these three shocks occur at a minute-to-minute frequency within a nurse's 12-hour shift, they isolate acute fluctuations in available decision time rather than slower-moving determinants of behavior such as fatigue, staffing levels, or cumulative workload. In balance tests presented in Appendix Figure A.1, I show that the three measures of time pressure are not associated with economically meaningful differences in observable pre-triage characteristics, supporting their interpretation as quasi-random variation in the time available for each decision.

3.2 Estimating Equations

I quantify how acute time pressure affects information acquisition and the triage decision itself using models that exploit within–nurse-by-shift variation in time pressure. The main specification represents time pressure using a set of discrete indicators for different pressure levels, allowing the relationship between time pressure and outcomes to be fully flexible and interpreted directly as changes relative to the lowest-pressure condition.

Information acquisition outcomes To characterize information compression prior to the triage decision, I examine how acute time pressure affects assessment intensity, measured by

minutes spent in triage, the number of diagnostic signals collected, and the length of the free-text triage note.

Let $\{1\{\text{TP}_{ijst} \in k\}\}_k$ denote mutually exclusive indicators for binned time-pressure states (e.g., number-waiting categories, next-patient-wait bins, or previous-triage-duration quintiles), with the lowest-pressure bin omitted. For each information-acquisition outcome, I estimate:

$$y_{ijst} = \alpha + \sum_k \delta_k \, 1\{\text{TP}_{ijst} \in k\} + X'_{ijst}\gamma + \mu_{js} + \lambda_{ht} + \varepsilon_{ijst}, \tag{1}$$

where y_{ijst} denotes an information-related outcome for patient i evaluated by nurse j during shift s at time t; TP_{ijst} indexes the level of time pressure; X_{ijst} includes patient demographics, arrival mode, chief-complaint fixed effects, and clinical covariates; μ_{js} are nurse-by-shift fixed effects; and λ_{ht} are hospital-by-hour fixed effects. The coefficients δ_k measure how assessment intensity changes relative to the lowest-pressure state, isolating the effect of acute time constraints on information collection and processing.

These estimates establish how time pressure constrains the information set available at triage, providing a mechanism through which time pressure can reshape triage decisions in the specifications that follow.

Triage priority and decision reweighting I next examine how acute time pressure reshapes triage decisions by altering the weight placed on coarse, readily observable patient characteristics. Using the same binned representation of time pressure, I estimate:

$$\begin{aligned}
\text{TopTriage}_{ijst} = \alpha &+ \sum_k \theta_k \, 1\{\text{TP}_{ijst} \in k\} + \beta^A \, \text{Age65}_i + \beta^G \, \text{Male}_i + \beta^R \, \text{White}_i \\
&+ \sum_k \beta^A_k \, 1\{\text{TP}_{ijst} \in k\} \times \text{Age65}_i + \sum_k \beta^G_k \, 1\{\text{TP}_{ijst} \in k\} \times \text{Male}_i \\
&+ \sum_k \beta^R_k \, 1\{\text{TP}_{ijst} \in k\} \times \text{White}_i \\
&+ X'_{ijst}\gamma + \mu_{js} + \lambda_{ht} + \varepsilon_{ijst},
\end{aligned} \tag{2}$$

where TopTriage_{ijst} is an indicator for receiving an ESI triage score of 1 or 2 (out of 5); Age65_i indicates age 65 or older; Male_i indicates male; and White_i indicates non-Hispanic

White. The coefficients θ_k capture the effect of time pressure on triage assignment for the omitted group, while the interaction terms β_k^A, β_k^G, and β_k^R trace how the weight placed on age, gender, and race in the triage decision varies with time pressure. Standard errors are clustered at the nurse level.[11]

3.3 Identifying Assumptions

Identification relies on the assumption that, conditional on rich observables and fixed effects, short-run fluctuations in time pressure are quasi-random and uncorrelated with unobserved determinants of patient severity or triage priority, including those that vary by patient demographics.

(1) Number of Patients Waiting When the waiting room is fuller, nurses experience greater pressure to move quickly through each assessment. Short-run fluctuations in the number of patients awaiting triage arise primarily from random patient arrivals. The identifying assumption is that these fluctuations are not systematically correlated with group-specific (e.g., gender, age, race) differences in the underlying severity of the current patient, conditional on rich patient observables such as chief complaint, vital signs, and arrival mode.

(2) Wait Time of Next Patient In Line Even holding queue length fixed, time pressure increases when the next patient in line has already been waiting longer. The identifying assumption is that, conditional on queue length and observables, variation in the next patient's wait time is unrelated to group-specific severity of the current patient.

(3) Duration of Previous Triage If the prior patient's triage took unusually long, the nurse may rush to "catch up" by shortening the next evaluation. Such variation reflects idiosyncratic differences in how long earlier assessments happened to take, rather than systematic differences in the current patient's condition. The identifying assumption is that variation in the prior triage's duration is not systematically related to group-specific severity of the current patient.

[11] Standard errors are clustered at the nurse level, since decision-making may be correlated across cases and shifts for the same nurse (Abadie et al., 2023).

Under these assumptions, the estimates that follow can be interpreted as reflecting the effects of acute time pressure on information compression and shifting decision weights.

4 Results

I now turn to the main results. I begin by showing how acute time pressure constrains the application of formal triage protocols by compressing information acquisition at the point of decision: nurses spend less time evaluating patients and collect fewer clinical signals when decision time is scarce. I then show that, as individualized assessment is truncated, triage decisions substitute away from fully protocolized evaluation toward simplified decision rules that place greater weight on coarse, readily observable patient characteristics, with the decision weight on age, gender, and race rising systematically under higher time pressure.

Next, I examine how these simplified decision rules are formed in practice. Focusing on race as a dimension along which prior exposure varies across decision-makers, I show that accumulated experience shapes how race is mapped into acuity assessments when formal information is limited. I then study how these experience-based rules are applied once activated, providing evidence that they spill over across cases and continue to influence decisions even when the cues that motivate their use are not relevant for the current patient. Finally, I examine the implications of this substitution from formal to informal decision rules, showing that racial gaps in downstream assessment and treatment widen under higher time pressure as early triage decisions propagate through subsequent stages of care.

4.1 Time Pressure Compresses Case-Specific Information

I begin by examining how acute time pressure affects the amount of case-specific information collected during triage. Because triage decisions rely on discretionary information acquisition under time constraints, increases in time pressure may influence the nurse's decision to gather individualized clinical signals. Using the three quasi-random sources of time-pressure variation described above, I test whether higher time pressure reduces assessment intensity, as measured by time spent in triage, the number of diagnostic signals collected, and the length of clinical notes.

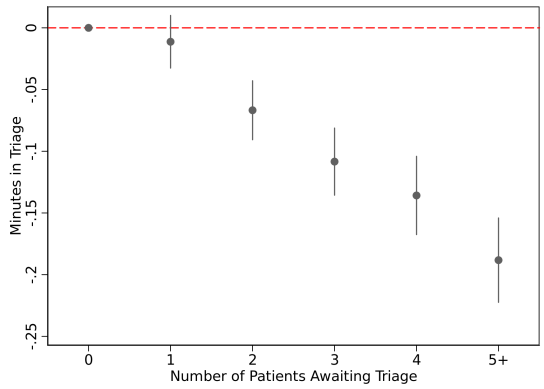
I first focus on time spent in triage as a direct measure of information acquisition intensity. To do so, I estimate Equation 1 with minutes spent in triage as the dependent variable. Across all three sources of quasi-random variation—the number of patients waiting, the waiting time of the next patient in line, and the duration of the immediately preceding triage—higher time pressure leads nurses to spend less time assessing the current patient. Figure 1 shows a clear and monotonic relationship in each case: relative to the lowest-pressure state, assessment duration falls steadily as time pressure increases. Moving from the reference bin to the highest-pressure bin reduces time spent in triage by approximately 4 to 6 percent. This change represents a material compression of an already brief, protocolized, and high-stakes clinical interaction.

Consistent with shorter assessment times, nurses also collect less individualized information when under time pressure. Figure 2 shows that the number of structured diagnostic fields completed declines as time pressure rises, and Appendix Figure A.2 shows a corresponding reduction in triage note length.[12] The absolute changes in these measures are modest, but the mapping from assessment duration or signal counts to decision-making is a priori unknown, and even marginal reductions in information can alter decision rules. Importantly, these patterns do not suggest that triage collapses under pressure; rather, they indicate that acute time constraints induce incremental losses of case-specific information. As shown next, these marginal reductions are sufficient to shift decision-making toward coarse, readily observable characteristics that are available prior to the assessment.

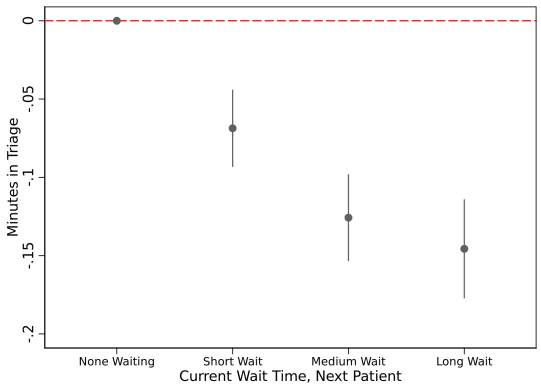
4.2 Time Pressure Shifts Decision Weights in Triage Decisions

I next examine how acute time pressure reshapes triage decisions by altering how formal triage protocols are applied in practice. Building on the evidence above that time pressure compresses case-specific information acquisition, this section studies whether nurses substitute away from fully individualized, protocolized assessment toward simplified decision rules

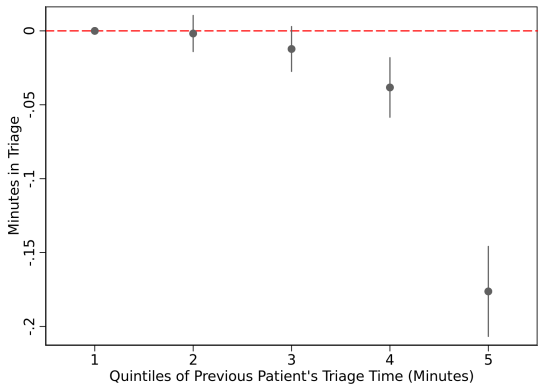
[12]Triage notes provide a useful proxy for cognitive effort and information processing. Prior work shows that when clinicians face higher workloads, they tend to write shorter, more formulaic notes, reflecting reduced cognitive engagement and processing (Hsu, Obermeyer and Tan, 2023). In this context, shorter notes likely indicate not only less documentation available to downstream providers, but also reduced processing of patient-specific information at the point of assessment.



(a) Patients Waiting



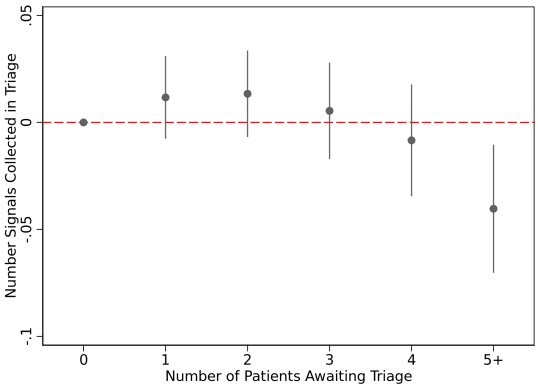
(b) Next Patient Wait



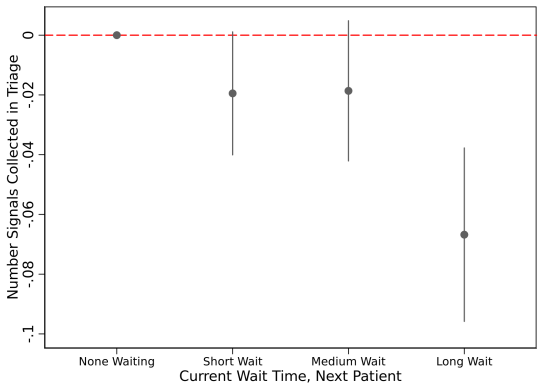
(c) Previous Triage Duration

Figure 1: Minutes in Triage Under Time Pressure

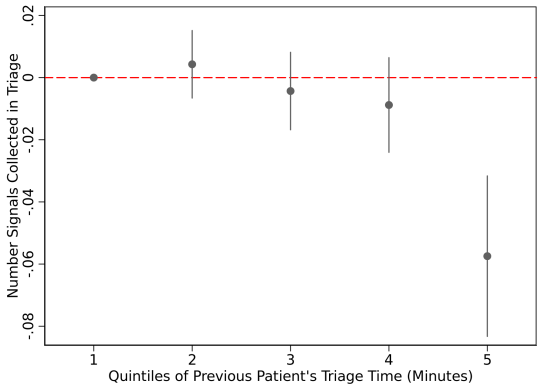
Note: Figure shows regression estimates of minutes spent in triage as a function of three quasi-random sources of time pressure. Panel (a) plots estimates by the number of patients awaiting triage, with 0 as the reference category (unconditional average triage time for patients with 0 people waiting: 4.2 minutes). Panel (b) plots estimates by the current wait time of the next patient in line, with "None Waiting" as the reference category. Panel (c) plots estimates by quintiles of the previous patient's triage time, with the shortest quintile as the reference category (unconditional average triage time for patients following a prior triage in the first quintile: 3.0 minutes). All regressions control for patient race (White vs. non-White), gender, a cubic in patient age, and an indicator for whether the patient was skipped over by someone else in the queue. The specification further absorbs nurse-by-12-hour-shift fixed effects, chief complaint fixed effects, arrival mode fixed effects, hospital-by-hour fixed effects, and the non-focal time pressure dimensions. Standard errors are clustered at the triage nurse level.



(a) Patients Waiting



(b) Next-Patient Wait



(c) Previous Triage Duration

Figure 2: Number of Diagnostic Signals Collected Under Time Pressure

Note: Figure shows regression estimates of the number of diagnostic signals collected during triage (out of 12) as a function of three quasi-random sources of time pressure. Panel (a) plots estimates by the number of patients awaiting triage, with 0 as the reference category. The unconditional average number of signals collected for patients with 0 people waiting is 9.0. Panel (b) plots estimates by the current wait time of the next patient in line, with "None Waiting" as the reference category. Panel (c) plots estimates by the duration of the previous patient's triage, with the shortest quintile as the reference category. The unconditional average number of signals collected for patients following a prior triage in the shortest quintile is 9.0. All regressions control for gender, a cubic in patient age, and an indicator for whether the patient was skipped over by someone else in the queue. The specification further absorbs nurse-by-12-hour-shift fixed effects, chief complaint fixed effects, arrival mode fixed effects, hospital-by-hour fixed effects, and the non-focal time dimensions. Standard errors are clustered at the triage nurse level.

that place greater weight on coarse, readily observable patient characteristics. I focus on age, gender, and race—features that are always observed at the moment of decision and require no additional information acquisition.

When time constraints limit the ability to collect and process case-specific signals, front-line decision-makers may rely more heavily on such coarse categories as inputs into decision-

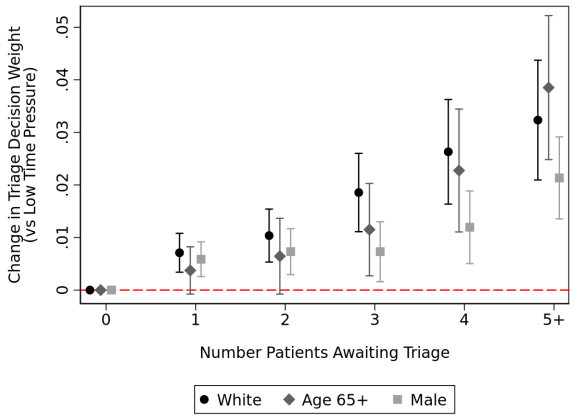
making, effectively reweighting how information enters the triage rule. The analysis therefore tests whether quasi-random, minute-to-minute increases in time pressure systematically shift triage decisions toward these readily available cues, consistent with endogenous substitution from formal to simplified decision rules under organizational constraints.

To do so, I estimate Equation 2, which interacts indicators for time pressure with patient age, gender, and race. The interaction coefficients capture how the contribution of each coarse characteristic to the triage decision varies with time pressure, holding fixed nurse-by-shift conditions, hospital-by-hour constraints, presenting complaint, arrival mode, and patient visit history. Figure 3 illustrates how these decision weights evolve across bins of time pressure. Across all three quasi-random sources of variation, the weight placed on these coarse characteristics increases monotonically as time pressure rises. Race exhibits the largest and most precisely estimated reweighting: relative to the lowest-pressure state, the White–non-White gap in the probability of receiving top triage priority increases by approximately 20–50 percent relative to its baseline level under the highest-pressure conditions, depending on the measure of pressure.

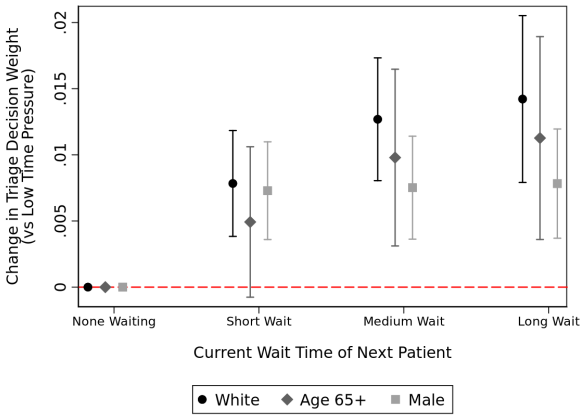
The comparable increase in decision weight placed on both age and race under time pressure raises a natural puzzle. Greater reliance on age when individualized information is compressed is consistent with triage protocols, which explicitly treat age as a coarse indicator of acute physiological risk. By contrast, triage guidelines are race-neutral: race does not enter the triage manual and plays no role in the protocol's mapping from observed symptoms to triage priority. The fact that race is reweighted alongside age under acute constraints therefore cannot be explained by formal triage rules. Section 4.3 investigates where race-based associations used under time pressure come from.

Alternative Explanations and Mechanisms A remaining concern is that the shifting decision weights I attribute to time pressure may instead reflect cumulative fatigue or declining cognitive capacity as nurses progress through their shifts. Two features of the design and results make this explanation unlikely.

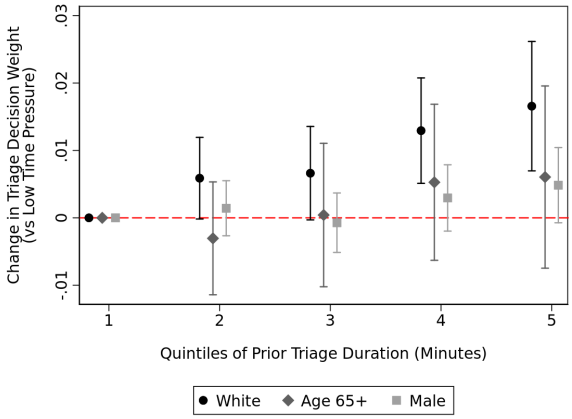
First, the sources of variation I exploit capture acute, high-frequency shocks to nurses' available decision time rather than gradual changes in energy or attention. Queue length,



(a) Patients Waiting



(b) Next-Patient Wait



(c) Previous Triage Duration

Figure 3: Coarse Decision Reweighting Under Time Pressure

Note: Figure reports regression estimates of how the implied weight placed on coarse, readily observable patient characteristics varies with acute time pressure. Each panel plots changes in the contribution of race (White), age (65+), and gender (Male) to the probability of receiving top triage priority, relative to the lowest-pressure state, across three quasi-random sources of time pressure. Panel (a) uses the number of patients awaiting triage, with zero patients waiting as the reference category; baseline gaps are 6.2 pp for race, 6.9 pp for age, and 2.1 pp for gender. Panel (b) uses the waiting time of the next patient in line, with "None Waiting" as the reference category. Panel (c) uses quintiles of the previous patient's triage duration, with the shortest quintile as the reference category; baseline gaps are 6.1 pp for race, 7.2 pp for age, and 2.3 pp for gender. All regressions include fixed effects for nurse-by-shift, hospital-by-hour, chief complaint, arrival mode, patient visit history, and the non-focal time pressure bins. Standard errors are clustered at the triage nurse level.

the waiting time of the next patient in line, and the duration of the immediately preceding triage fluctuate from minute to minute due to random patient arrivals and idiosyncratic variation in triage duration. By construction, these high-frequency changes are unlikely to

be systematically related to the slow-moving accumulation of fatigue over a 12-hour shift.

Second, I present two sets of empirical results that are inconsistent with a fatigue-based interpretation. If cumulative fatigue were driving the widening disparities, the racial gap should increase systematically later in the shift. Instead, Appendix Table A.1 shows that the estimated effects of time pressure on triage decisions are broadly similar across shift halves. For example, high-time pressure widens the White–non-White gap across all three time-pressure measures, while interactions with second-half indicators are statistically insignificant for the queue-length and previous-duration measures and small and unstable for the next-patient-wait measure. Appendix Table A.2 corroborates this conclusion: the estimated time-pressure effects remain stable when I flexibly control for within-shift progress using quintile fixed effects and their interactions with White, which absorb average changes in fatigue, pacing, and triage stringency as shifts progress. Taken together, these results indicate that the shifting decision weights I document are driven by acute time constraints rather than by the gradual buildup of fatigue or cognitive depletion over the course of a shift.

4.3 Experience-Based Associations in Triage Decisions

Having shown that time pressure induces substitution from fully protocolized assessment toward simplified decision rules, I next examine how the content of these rules is formed in practice. In organizational settings where frontline workers repeatedly evaluate similar cases, experience can shape how coarse characteristics are mapped into decision-relevant assessments when formal information is limited. When exposure to particular types of cases differs systematically across decision-makers, the simplified rules they rely on under constraint may reflect the patterns most frequently encountered in their recent experience.

Consistent with this perspective, a theoretical and experimental literature emphasizes that repeated exposure can generate coarse category-based associations that are subsequently deployed when decision-makers face binding constraints on information acquisition (Mullainathan, Schwartzstein and Shleifer, 2008; Bordalo et al., 2016). This section examines whether recent experience shapes how race enters triage decisions under time pressure, providing evidence on how informal decision rules emerge endogenously from organizational exposure rather than from fixed beliefs or preferences.

In the ED, one salient and widely recognized patient category is the "frequent flier" — patients with four or more visits in the prior 12 months —who are often perceived as less urgent (Grover and Close, 2009; LaCalle, Rabin and Genes, 2013; Krieg et al., 2016). Importantly, frequent flier status is unevenly distributed across racial groups. In my data, 22.74% of visits by non-White patients involve frequent fliers, compared with 11.57% among White patients. Consistent with this, among all frequent flier encounters, 73.34% involve non-White patients. As a result, repeated exposure to frequent fliers disproportionately exposes nurses to non-White patients in lower-acuity contexts, making race a natural candidate for encoding associations about urgency that are formed through experience.

If nurses' decisions draw on associations between race and clinical severity that are shaped through experience with frequent fliers, then prior exposure to non-White frequent fliers should be associated with stronger race-based gradients in triage decisions, while prior exposure to White frequent fliers should be associated with weaker ones. To examine this possibility, I analyze how the within–nurse-by-shift, White–non-White gap in top triage priority varies with a nurse's recent exposure to frequent fliers. I measure exposure as the share of a nurse's total patient encounters over the prior 90 days that involved frequent fliers, computed before the start of the focal shift. Because this exposure measure is fixed within a shift and varies only across shifts of the same nurse or across different nurses, identification comes from comparing within–nurse-by-shift racial gaps across shifts with different levels of prior exposure. All specifications exclude patients who are themselves current frequent fliers, ensuring that the exposure measure is not mechanically related to the racial composition or acuity of the current patient.

Figure 4 plots the estimated White–non-White gap in top triage assignment across quintiles of prior exposure to non-White and White frequent fliers. Two patterns stand out. First, the racial gap increases sharply with exposure to non-White frequent fliers: relative to the lowest-exposure quintile, the gap in the highest quintile is larger by approximately 8.4 percentage points. Second, exposure to White frequent fliers is associated with an attenuation of the racial gap: nurse-shifts in the highest exposure quintile exhibit a gap that is roughly 2.2 percentage points smaller than those in the lowest quintile. These relationships are estimated within nurse-by-shift and condition on chief complaint, arrival mode, visit his-

tory, and hospital-by-hour crowding conditions, helping to distinguish longer-run exposure patterns from differences in the composition of the current shift.

Taken together, these results indicate that repeated exposure to particular types of cases shapes the content of the simplified decision rules that are applied under time pressure. In this setting, frequent ED users are disproportionately non-White and are commonly treated as lower acuity, so greater recent exposure to non-White frequent users is associated with larger racial gradients in triage priority, while greater exposure to White frequent users attenuates these gradients. These patterns suggest a concrete organizational mechanism through which experience-based learning maps coarse characteristics into acuity assessments when formal, case-specific information is limited. Rather than reflecting fixed beliefs or preferences, the role of race in triage decisions appears to emerge endogenously from the distribution of cases that frontline workers encounter in their recent workflow.

4.4 Persistence and Spillovers of Simplified Decision Rules

A central question for organizational decision-making is not only how informal decision rules are formed, but how they are applied once activated. In settings characterized by rapid, sequential decisions, simplified rules that are triggered by local conditions may persist beyond the situations that motivate their use, influencing subsequent decisions with limited case-by-case selectivity. Such spillovers are a common feature of organizational routines, particularly when decisions must be made quickly and attention is constrained.

Building on the evidence above that experience shapes how race enters simplified triage rules under time pressure, this section examines whether these rules spill over across patients within the triage queue. Specifically, I test whether encounters with frequent ED users—whose characteristics contain no information about the acuity of the next patient—activate race-based assessments that continue to influence subsequent triage decisions. If nurses rely on simplified, experience-based rules when formal, case-specific information is limited, then these rules may persist across decisions even when they are not informative for the current case.

Table 2 reports estimates from regressions with an indicator for receiving top triage priority as the dependent variable. The specification includes an indicator for non-Hispanic White

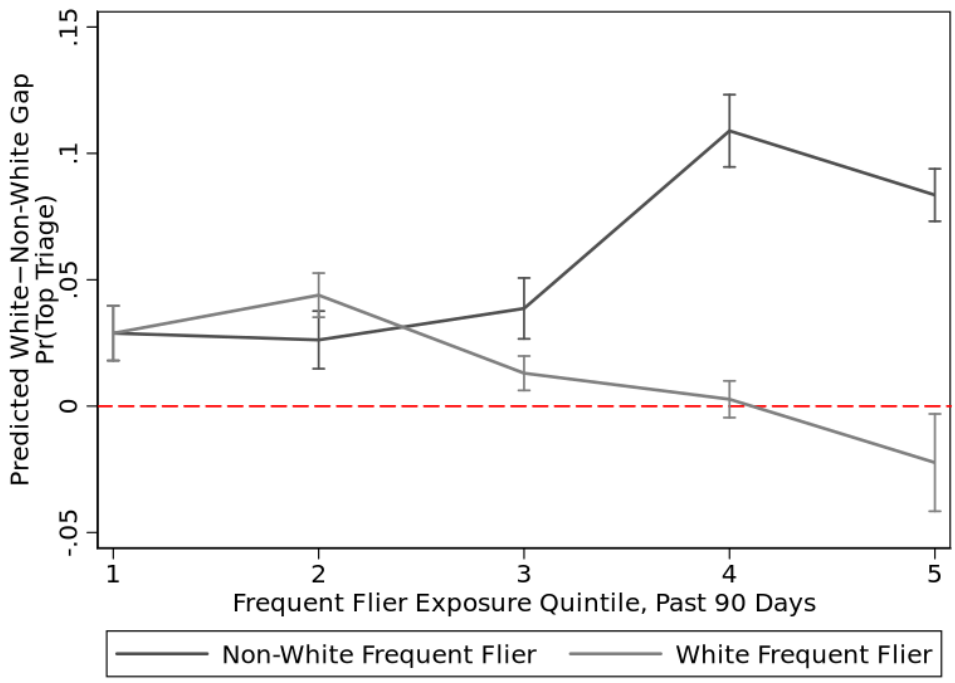


Figure 4: Racial Disparities in Top Triage Probability by Frequent Flier Exposure

Note: Figure shows regression estimates of the White–non-White differences in the probability of top triage by pooled quintiles of a nurse's recent exposure (prior 90 days, frozen at shift start) to non-White frequent fliers and White frequent fliers. Q1 (lowest exposure) is the reference category. The Q1 predicted gap is 2.89 pp. Frequent fliers are defined as patients with four or more visits in the prior 12 months. The specification includes both sets of exposure-by-White interactions simultaneously, controls for gender and a cubic in age, and absorbs nurse-by-12-hour-shift, chief complaint, and arrival mode fixed effects. The analysis sample excludes current frequent fliers and requires ≥ 30 prior visits to compute exposure. Quintile mean exposure shares: non-White FF — Q1 1.3, Q2 3.1, Q3 5.0, Q4 11.8, Q5 20.9; White FF — Q1 1.2, Q2 3.2, Q3 5.3, Q4 8.7, Q5 18.0. 95 percent confidence intervals shown; standard errors clustered at the triage nurse level.

patients and interactions between this indicator and whether the immediately preceding patient was a frequent flier, distinguishing between cases in which the preceding frequent flier was White versus non-White. The coefficient on the White indicator captures the baseline White–non-White gap in top triage assignment, while the interaction terms capture how this gap changes following encounters with White and non-White frequent fliers.[13]

The results indicate that race-correlated decision rules persist across sequential triage decisions, even when they are not informative about the acuity of the current patient. Following an encounter with a non-White frequent ED user, the White–non-White gap in top

[13]The specification includes an indicator for whether the current patient is a frequent flier, ensuring that estimated spillovers are not mechanically driven by the frequent flier status of the focal patient.

Table 2: Spillovers of Simplified Decision Rules

	Pr(Top Triage)
White	0.060***
	(0.003)
White × After Non-White FF	0.014***
	(0.003)
White × After White FF	-0.005*
	(0.003)
Observations	1,800,398
Nurses	1,661
R^2	0.386
Nurse–Shift FE	Yes
Hospital–Hour FE	Yes
Ambulance FE	Yes
Chief Complaint FE	Yes
Visit Number FE	Yes

Note: $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Table reports estimates from regressions of an indicator for receiving top triage priority (ESI 1–2) on patient race and interactions between the race and the frequent flier status of the immediately preceding patient. The coefficient on *White* captures the baseline White–non-White gap in top triage assignment for patients who do not follow a frequent flier. The interaction terms *White* × *After Non-White FF* and *White* × *After White FF* capture how this racial gap changes following an encounter with a non-White or White frequent flier, respectively. All specifications include nurse-by-shift fixed effects, hospital-hour fixed effects, chief complaint fixed effects, arrival mode, visit history controls, and an indicator for whether the current patient is a frequent flier. Standard errors are clustered at the triage nurse level.

triage assignment increases by approximately 23% relative to baseline. By contrast, following an encounter with a White frequent ED user, the racial gap attenuates by roughly 9% ($p < 0.10$). Because the race and clinical characteristics of the current patient are orthogonal to whether the preceding patient was a frequent ED user, these sequential shifts cannot reflect updated information about patient acuity. Instead, they indicate that experience-based decision rules—once activated by recent encounters—carry over across cases, influencing subsequent decisions with limited case-by-case selectivity.

Taken together, the evidence in this section shows that race enters triage decisions through the formation and activation of simplified decision rules shaped by organizational experience and immediate context. Longer-run exposure to racially patterned patient types is

associated with systematic differences in how race is incorporated into triage decisions across shifts, while encounters with frequent ED users within a shift activate these experience-based rules in real time. These patterns indicate that race influences triage decisions not because it is inherently informative about clinical severity, but because repeated exposure shapes how coarse characteristics are mapped into acuity assessments when formal, case-specific information is limited. Having established how these simplified decision rules are formed and activated, I next examine how they propagate through subsequent stages of patient evaluation and treatment.

5 Downstream Implications of Simplified Decision Rules

Triage decisions constitute an early input that structures subsequent workflows during an ED visit. By assigning priority at intake, triage assessments shape how attention, information, and treatment resources are allocated across later stages of care. When early decisions rely on simplified decision rules rather than fully individualized assessment, these rules can influence not only the initial assignment of priority but also how subsequent decisions unfold.

Building on the evidence above that acute time pressure compresses information acquisition and induces substitution toward simplified triage rules, this section examines whether time pressure at the point of triage is associated with systematic differences in downstream evaluation and treatment. The analysis asks whether early constraints faced by frontline decision-makers propagate through organizational processes, generating predictable differences in care even after patients move beyond the initial decision stage.[14]

I first examine delays in initial evaluation. Table 3 shows that among higher-acuity encounters (ESI 1–3), high nurse time pressure is associated with longer wait times for non-White patients and a widening White–non-White gap in time to first physician contact. Under high pressure, delays increase by several minutes for non-White patients, while changes for White patients are smaller. These effects are concentrated in urgent cases treated in the main ED and are absent for lower-acuity visits routed to Fast Track areas, where queueing

[14]Downstream analyses do not condition on triage priority or related assessments conducted by the nurse, as these variables are endogenous to time pressure and constitute part of the mechanism through which early decision rules shape subsequent stages of care (Angrist and Pischke, 2009).

and staffing operate differently.[15]

Table 3: Time Pressure and Racial Disparities in Wait Times

	(1) ESI 1–3 (Urgent)	(2) ESI 4–5 (Fast Track Eligible)
High Time Pressure (TP)	3.149***	0.566*
	(0.276)	(0.291)
Non-Hispanic White (Low TP)	-2.661***	-1.331***
	(0.292)	(0.333)
High TP × White	-3.315***	0.045
	(0.340)	(0.445)
Observations	932,381	299,367
Nurses	1,495	1,149
R-sq	0.506	0.628
Nurse-Shift FE	Yes	Yes
Hosp-Hour FE	Yes	Yes
Ambulance FE	Yes	Yes
Chief Complaint FE	Yes	Yes
Visit Number FE	Yes	Yes

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable is minutes from triage to being seen by a physician. Average wait time for non-White patients under low time pressure is 56.04 minutes in Column (1) and 51.24 minutes in Column (2). "High Time Pressure (High TP)" equals 1 if the visit falls in the top two bins of *any* of the three time-pressure proxies (number of patients waiting, next-patient wait time, and previous triage duration; see text). Controls include a male indicator and a cubic in age. Fixed effects: nurse×shift, hospital-hour, arrival mode, chief complaint, and visit number. Standard errors (in parentheses) are clustered by triage nurse. Column (1) restricts to ESI 1–3 (urgent); Column (2) restricts to ESI 4–5 (low acuity/Fast Track).

I next examine delays relative to established clinical benchmarks for time-sensitive conditions. Table 4 shows that high time pressure is associated with longer delays beyond recommended targets for stroke, sepsis, and heart attack presentations, with larger delays for non-White patients. These patterns indicate that time pressure at triage is associated not only with short-lived waiting differences, but also with clinically meaningful delays in evaluation for the most time-sensitive conditions. Clinical benchmarks for stroke suggest that the increase in the racial delay gap under high time pressure, relative to low time pressure, corresponds to a meaningful increase in neurological damage and an approximately 0.4 pp increase in permanent disability risk for affected patients (Saver, 2006).

[15]Fast Track areas are designed to reduce congestion in the main ED by diverting and treating low acuity patients quickly. Because wait times are determined by different staffing models, patient flows, and queueing rules in the main ED and Fast Track, pooling the two settings would mix fundamentally different care processes.

Table 4: Time Pressure and Racial Disparities in Delays for Stroke, Sepsis, and Heart Attack

	(1) All Time-Sensitive (ESI 1–3)	(2) Stroke	(3) Sepsis	(4) Heart Attack
High Time Pressure (TP)	3.941***	4.666***	5.686***	2.301**
	(0.586)	(0.944)	(1.513)	(1.072)
Non-Hispanic White (Low TP)	-3.209***	-2.977***	-3.735***	-3.182***
	(0.562)	(0.823)	(1.327)	(1.123)
High TP × White	-4.570***	-4.433***	-5.117**	-1.530
	(0.828)	(1.192)	(2.115)	(1.707)
Observations	168,393	66,163	20,302	50,251
Nurses	1,195	1,053	789	913
R-sq	0.600	0.623	0.649	0.677
Nurse–Shift FE	Yes	Yes	Yes	Yes
Hospital–Hour FE	Yes	Yes	Yes	Yes
Ambulance FE	Yes	Yes	Yes	Yes
Chief Complaint FE	Yes	No	No	No
Visit Number FE	Yes	Yes	Yes	Yes

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable is the number of minutes by which the evaluation exceeded the clinical target for time-sensitive conditions (stroke, sepsis, and heart attack). Clinical targets are defined as 20 minutes for stroke (door-to-CT), 30 minutes for sepsis (door-to-physician evaluation), and 10 minutes for heart attack (door-to-ECG). Average delay for non-White patients under low time pressure is 43.72 minutes in the pooled sample, 36.87 minutes for stroke, 27.03 minutes for sepsis, and 56.50 minutes for heart attack. "High Time Pressure (High TP)" equals one if the triage nurse was in the top two bins of any of the three time-pressure proxies (number of patients waiting, next-patient wait time, or previous triage duration; see text). Controls include a male indicator and a cubic in age. Fixed effects: nurse×shift, hospital-hour, arrival mode, and visit number. Column (1) additionally includes indicators for the specific time-sensitive condition. Standard errors (in parentheses) clustered by triage nurse. Column (1) pools all time-sensitive visits (ESI 1–3); Columns (2)–(4) restrict the sample to stroke, sepsis, and heart attack presentations, respectively. Patients seen ahead of clinical targets are treated as having zero delay.

Finally, I examine downstream diagnostic effort, treatment, and hospital admission. Appendix Tables A.3 and A.4 show that under high time pressure, non-White patients receive fewer diagnostic assessments and are less likely to receive procedures, pain medication, or be admitted to the hospital, while corresponding changes for White patients are smaller or absent. Because downstream clinicians observe triage priority and nurse documentation at the start of evaluation, these results are consistent with early differences in triage assessment propagating through subsequent stages of care.[16] Because these downstream assessments play a central role in confirming or correcting the initial triage signal, these patterns indicate that time pressure is associated with increasingly divergent information environments for White and non-White patients as care unfolds.

[16] Also see Singh and Venkataramani (2024) for evidence that system-wide capacity strain is associated with racial gaps in mortality; the present paper sheds light on one micro-level foundation through which such disparities may arise.

Taken together, these findings suggest that acute time pressure at triage is associated with persistent differences in how attention, information, and treatment are allocated downstream, consistent with the mechanism of information compression and decision reweighting documented earlier.

6 Conclusion

This paper shows that acute time pressure can fundamentally reshape how decisions are made inside organizations. Using high-frequency administrative data from emergency department triage, I exploit short-run, plausibly exogenous fluctuations in available decision time to study how frontline workers apply formal decision rules in practice. I find that when time is scarce, nurses compress individualized information acquisition and substitute away from formally prescribed decision rules toward simplified, experience-based rules that place greater weight on coarse, readily observable characteristics.

These shifts reflect changes in decision rules rather than increases in noise around a fixed rule. Under time pressure, nurses spend less time evaluating patients, collect fewer diagnostic signals, and systematically reweight which inputs enter triage decisions. The simplified rules that emerge are shaped by accumulated experience and activated by immediate context, providing field evidence on how informal decision rules form and operate alongside formal organizational procedures. Once activated, these rules can persist across cases, even when the cues that motivate their use are not relevant to the current decision.

Because triage assessments structure subsequent clinical workflows, deviations at this initial stage propagate through downstream decisions. Patients evaluated under higher time pressure experience longer delays, reduced diagnostic assessment, and lower treatment intensity, with effects differing systematically across groups. These patterns illustrate how localized frictions faced by individual decision-makers can generate persistent and predictable differences in outcomes, even when formal rules, incentives, and personnel remain unchanged.

More broadly, the findings contribute to organizational economics by identifying time pressure as a key friction that limits the effectiveness of formal rules. While existing work emphasizes agency problems, incentive misalignment, and limited enforcement as sources of

divergence between formal organizational design and realized behavior, this paper highlights a complementary mechanism: acute constraints on decision time induce endogenous substitution from formal rules toward informal decision rules. In this sense, deviations from formal procedures need not reflect changes in preferences, incentives, or compliance, but can arise mechanically from constraints on information acquisition at the moment of choice.

The results also speak to debates on discrimination and unequal treatment by showing how disparities can emerge from the interaction of organizational constraints and experience-based decision rules. When information acquisition is truncated and simplified rules are applied broadly under pressure, group differences can arise and intensify even in the absence of explicit intent. Traditional distinctions between taste-based and statistical discrimination therefore capture only part of the forces shaping decisions in time-pressured organizational settings.

Taken together, the findings underscore the importance of organizational and job design in shaping decision-making. Across settings in which frontline workers must act quickly under formal protocols, short-run workload fluctuations can meaningfully alter how closely decisions reflect case-specific information. Organizational features that govern staffing flexibility, task allocation, queue management, and exposure to demand shocks may therefore influence not only efficiency, but also the degree to which formal rules discipline behavior in practice. More generally, the paper highlights the value of incorporating acute constraints on time and attention into economic models of organizations and decision-making.

References

Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey M Wooldridge. 2023. "When Should You Adjust Standard Errors For Clustering?" *The Quarterly Journal of Economics*, 138(1): 1–35.

Akepanidtaworn, Klakow, Rick di Mascio, Alex Imas, and Lawrence DW Schmidt. 2023. "Selling Fast and Buying Slow: Heuristics and Trading Performance of Institutional Investors." *The Journal of Finance*, 78(6): 3055–3098.

Alsan, Marcella, Katherine Ianni, and Graeme Peterson. 2024. "Racial Health Disparities in The United States." *Oxford Review of Economic Policy*, 40(3): 498–517.

Angrist, Joshua D, and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton university press.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul. 2007. "Incentives for Managers and Inequality among Workers: Evidence from a Firm-Level Experiment." *The Quarterly Journal of Economics*, 122(2): 729–773.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny. 1998. "A Model of Investor Sentiment." *Journal of Financial Economics*, 49(3): 307–343.

Bertrand, Marianne, and Esther Duflo. 2017. "Chapter 8 – Field Experiments on Discrimination." *Handbook of Economic Field Experiments*, 1: 309–393. eds. Abhijit Vinayak Banerjee and Esther Duflo, North-Holland.

Blonz, Joshua A. 2023. "The Costs of Misaligned Incentives: Energy Inefficiency and The Principal-Agent Problem." *American Economic Journal: Economic Policy*, 15(3): 286–321.

Bohren, J. Aislinn, Kareem Haggag, Alex Imas, and Devin G. Pope. 2023. "Inaccurate Statistical Discrimination: An Identification Problem." *The Review of Economics and Statistics.*

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer. 2016. "Stereotypes." *The Quarterly Journal of Economics*, 131(4): 1753–1794.

Bordalo, Pedro, Nicola Gennaioli, Giacomo Lanzani, and Andrei Shleifer. 2025. "A Cognitive Theory of Reasoning and Choice." NBER Working Paper 33466.

Chen, Daniel L., Tobias J. Moskowitz, and Kelly Shue. 2016. "Decision Making Under the Gambler's Fallacy: Evidence from Asylum Judges, Loan Officers, and Baseball Umpires." *The Quarterly Journal of Economics*, 131(3): 1181–1242.

Freedman, Seth, Ezra Golberstein, Tsan-Yao Huang, David J Satin, and Laura Barrie Smith. 2021. "Docs With Their Eyes on The Clock? The Effect of Time Pressures on Primary Care Productivity." *Journal of Health Economics*, 77.

Grover, Casey A., and Reb JH Close. 2009. "Frequent Users of The Emergency Department: Risky Business." *Western Journal of Emergency Medicine*, 10(3).

Handel, Ben, Louis Heizlsperger, Jonas Knecht, Jonathan Kolstad, Ulrike Malmendier, and Filip Matejka. 2025. "Thinking Versus Doing: Cognitive Capacity, Decision Making and Medical Diagnosis." *Working Paper*.

Hsu, Chao-Chun, Ziad Obermeyer, and Chenhao Tan. 2023. "Clinical Notes Reveal Physician Fatigue." *Working Paper*.

Imas, Alex, Michael A. Kuhn, and Vera Mironova. 2022. "Waiting to Choose: The Role of Deliberation in Intertemporal Choice." *American Economic Journal: Microeconomics*, 14(3): 414–440.

Jin, Lawrence, Rui Tang, Han Ye, Junjian Yi, and Songfa Zhong. 2020. "Time Dependency in Physician Decision-Making." *AEA Papers and Proceedings*, 110: 284–88.

Jin, Lawrence, Rui Tang, Han Ye, Junjian Yi, and Songfa Zhong. 2023. "Path Dependency in Physician Decisions." *The Review of Economic Studies*, 91(5): 2916–2953.

Kessler, Judd B., Hannu Kivimaki, Ashley Litwin, and Muriel Niederle. 2025. "Please Take A Minute: How Prosocial Choices Change With Deliberation." *Working Paper*.

Kocher, Martin G, Julius Pahlke, and Stefan T Trautmann. 2013. "Tempus Fugit: Time Pressure In Risky Decisions." *Management Science*, 59(10): 2380–2391.

Krieg, Cynthia, Catherine Hudon, Maud-Christine Chouinard, and Isabelle Dufour. 2016. "Individual Predictors of Frequent Emergency Department Use: A Scoping Review." *BMC Health Services Research*, 16(1).

LaCalle, Eduardo J., Elaine J. Rabin, and Nicholas G. Genes. 2013. "High-Frequency Users of Emergency Department Care." *The Journal of Emergency Medicine*, 44(6): 1167–1173.

Lang, Kevin, and Ariella Kahn-Lang Spitzer. 2020. "Race Discrimination: An Economic Perspective." *Journal of Economic Perspectives*, 34(2): 68–89.

Lawson, M. Asher, Richard P. Larrick, and Jack B. Soll. 2020. "Comparing Fast Thinking and Slow Thinking: The Relative Benefits of Interventions, Individual Differences, and Inferential Rules." *Judgment and Decision Making*, 15(5): 660–684.

Litwin, Ashley, and Corinne Low. 2025. "Measuring Discrimination with Experiments." *Handbook of Experimental Methods in the Social Sciences*. ed. Alex Rees-Jones, Edward Elgar Publishing Ltd.

Malmendier, Ulrike, and Stefan Nagel. 2011. "Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?" *The Quarterly Journal of Economics*, 126(1): 373–416.

Malmendier, Ulrike, and Stefan Nagel. 2016. "Learning from Inflation Experiences." *The Quarterly Journal of Economics*, 131(1): 53–87.

Mullainathan, Sendhil, Joshua Schwartzstein, and Andrei Shleifer. 2008. "Coarse Thinking and Persuasion." *The Quarterly Journal of Economics*, 123(2): 577–619.

Olken, Benjamin A. 2007. "Monitoring Corruption: Evidence From A Field Experiment In Indonesia." *Journal of Political Economy*, 115(2): 200–249.

Radbruch, Jonas, and Amelie Schiprowski. 2024. "Interview Sequences and the Formation of Subjective Assessments." *The Review of Economic Studies*, 92(2): 1226–1256.

Rose, Evan K. 2023. "A Constructivist Perspective on Empirical Discrimination Research." *Journal of Economic Literature*, 61(3): 906–923.

Saver, Jeffrey L. 2006. "Time is Brain — Quantified." *Stroke*, 37(1): 263–266.

Silver, David. 2021. "Haste or Waste? Peer Pressure and Productivity in the Emergency Department." *The Review of Economic Studies*, 88(3 (320)): 1385–1417.

Singh, Manasvini. 2021. "Heuristics In The Delivery Room." *Science*, 374(6565): 324–329.

Singh, Manasvini, and Atheendar Venkataramani. 2024. "Rationing by Race." *NBER Working Paper 30380*.

Worth, Murtis, Leslie L. Davis, Debra C. Wallace, Robin Bartlett, and Debbie Travers. 2019. "Are Emergency Departments In The United States Following Recommendations by The Emergency Severity Index To Promote Quality Triage and Reliability?" *Journal of Emergency Nursing*, 45(6): 677–684.

A Appendix

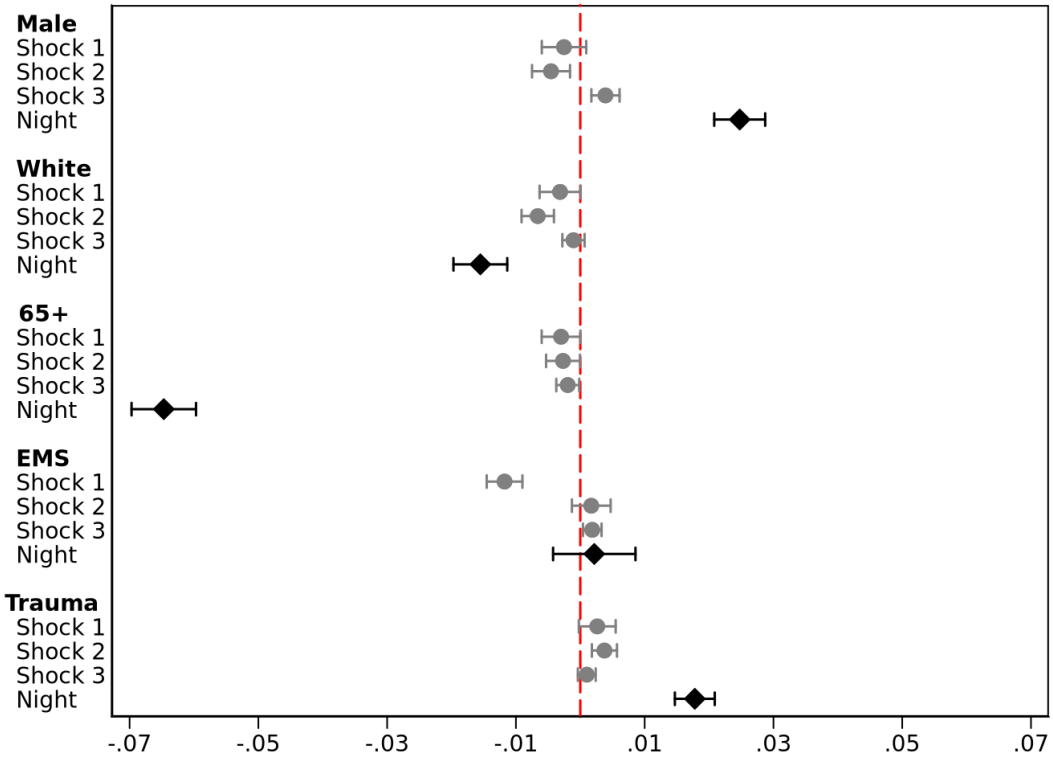
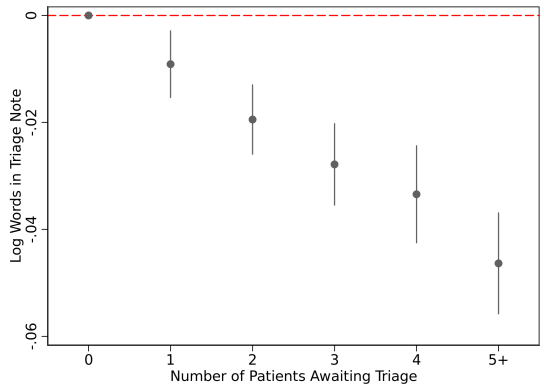
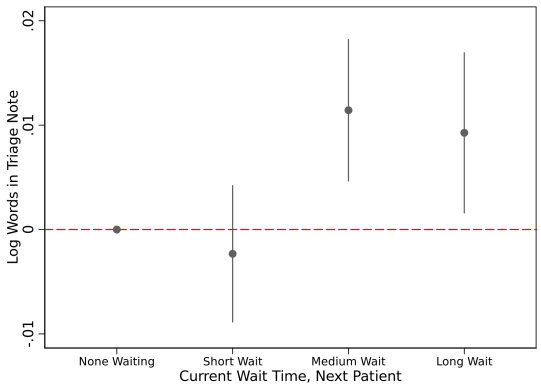


Figure A.1: Balance on Pre-Triage Characteristics

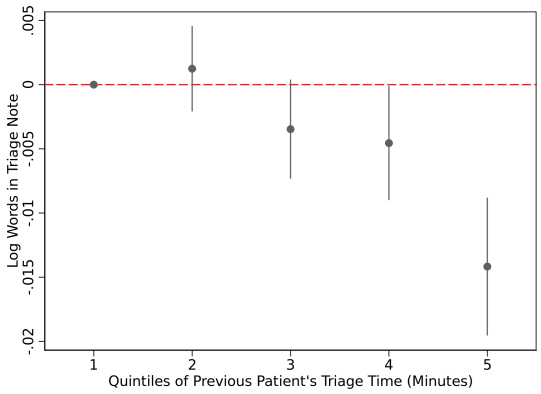
Note: Figure shows estimates from separate regressions of pre-determined patient characteristics on each indicator shown. Points denote point estimates and horizontal bars represent 95% confidence intervals. "Shock 1" is an indicator equal to one for patient visits in the top two bins of the number of patients waiting; "Shock 2" is an indicator for visits in the top two bins of the next patient's wait time; and "Shock 3" is an indicator for visits in the top two bins of the duration of the previous patient's triage. Nighttime arrival (7pm–7am) is included as an endogenous benchmark. Regressions with time-pressure indicators include fixed effects for nurse-shift, visit number, hospital-by-hour, and binned non-focal time-pressure shocks. Regressions for nighttime arrival include fixed effects for nurse-shift and visit number. Standard errors are clustered at the triage nurse level.



(a) Patients Waiting



(b) Next Patient Wait



(c) Previous Triage Duration

Figure A.2: Triage Note Length Under Time Pressure

Note: Figure shows regression estimates of the log number of words in triage notes as a function of time pressure, using three quasi-random sources of time pressure. Panel (a) plots estimates by the number of patients awaiting triage, with 0 as the reference category. The unconditional average note length for patients with 0 people waiting is 3.3 log words. Panel (b) plots estimates by the current wait time of the next patient in line, with "None Waiting" as the reference category. Panel (c) plots estimates by the duration of the previous patient's triage, with the shortest quintile as the reference category. The unconditional average note length for patients following a prior triage in the shortest quintile is 3.2 log words. All regressions control for race (White vs. non-White), gender, a cubic in patient age, and an indicator for whether the patient was skipped over by someone else in the queue. The specification further absorbs nurse-by-12-hour-shift fixed effects, chief complaint fixed effects, arrival mode fixed effects, hospital-by-hour fixed effects, visit number fixed effects, and the non-focal time dimensions. Standard errors are clustered at the triage nurse level.

Table A.1: Time Pressure vs. Cumulative Fatigue (First vs. Last Half of Shift)

	(1) Num Waiting	(2) Next-Patient Wait	(3) Prev Duration
High TP × White	0.021***		
	(0.005)		
High TP × White × First Half	-0.002		
	(0.005)		
High TP × White		0.014***	
		(0.003)	
High TP × White × First Half		-0.005*	
		(0.003)	
High TP × White			0.006*
			(0.003)
High TP × White × First Half			0.005
			(0.003)
Observations	1,487,511	1,487,511	1,337,858
Nurses	1,519	1,519	1,484
R-sq	0.363	0.363	0.374
Nurse-Shift FE	Yes	Yes	Yes
Hosp-Hour FE	Yes	Yes	Yes
Ambulance FE	Yes	Yes	Yes
Chief complaint FE	Yes	Yes	Yes
Visit number FE	Yes	Yes	Yes
Num Waiting FE	Yes	Yes	Yes
Next-Patient Wait FE	Yes	Yes	Yes
Prev Duration FE	Yes	Yes	Yes

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable is an indicator for receiving top triage priority. "High Time Pressure (High TP)" equals one for observations in the top two bins of each time pressure proxy. Each column reports the estimated change in the White–non-White gap under High versus Low time pressure ("High TP × White"), and the difference in that change between the first and second halves of the shift ("High TP × White × First Half"). All specifications include a male indicator, a cubic in age, and absorb fixed effects for nurse×shift, chief complaint, arrival mode, hosp-hour of arrival, visit number, and the binned versions of all time pressure proxies. Standard errors are clustered by triage nurse.

Table A.2: Time Pressure vs. Cumulative Fatigue (Shift Progress Quintiles)

	(1) Num Waiting	(2) Next-Patient Wait	(3) Prev Duration
White × High TP	0.021***		
	(0.004)		
White × High TP		0.012***	
		(0.002)	
White × High TP			0.008***
			(0.003)
Observations	1,487,511	1,487,511	1,337,858
Nurses	1,519	1,519	1,484
R-sq	0.363	0.363	0.374
Nurse-Shift FE	Yes	Yes	Yes
Hosp-Hour FE	Yes	Yes	Yes
Ambulance FE	Yes	Yes	Yes
Chief complaint FE	Yes	Yes	Yes
Visit number FE	Yes	Yes	Yes
Num Waiting FE	Yes	Yes	Yes
Next-Patient Wait FE	Yes	Yes	Yes
Prev Duration FE	Yes	Yes	Yes

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable is an indicator for receiving top triage priority. "High Time Pressure (High TP)" equals one for observations in the top two bins of each time pressure proxy. Each column reports the estimated change in the White–non-White gap under High versus Low time pressure ("High TP × White"). All specifications include a male indicator, a cubic in age, and absorb fixed effects for nurse×shift, chief complaint, arrival mode, hosp-hour of arrival, visit number, and the binned versions of all time pressure proxies. To flexibly account for within-shift drift in triage behavior—such as gradual changes in fatigue, pacing, or baseline decision stringency—each regression also includes quintile fixed effects of shift progress and their interactions with White. Standard errors are clustered by triage nurse.

Table A.3: Time Pressure and Racial Disparities in Downstream Information Acquisition

	Imaging		Labs		Vitals		Pain	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Any	Number	Any	Number	Any	Number	Any	Number
High Time Pressure (TP)	-0.001	-0.013**	-0.002**	-0.747***	-0.002**	-0.106***	-0.004**	-0.053*
	(0.001)	(0.006)	(0.001)	(0.169)	(0.001)	(0.026)	(0.002)	(0.031)
Non-Hispanic White (Low TP)	0.021***	0.089***	0.004***	2.775***	0.017***	0.646***	0.022***	0.644***
	(0.002)	(0.009)	(0.001)	(0.337)	(0.001)	(0.051)	(0.002)	(0.049)
High TP × White	0.004**	0.027***	0.003***	1.724***	0.002	0.237***	0.013***	0.114**
	(0.002)	(0.008)	(0.001)	(0.255)	(0.002)	(0.039)	(0.002)	(0.045)
Observations	1,337,499	951,099	1,334,655	1,164,606	1,262,436	1,019,105	1,260,625	620,912
Nurses	1,629	1,593	1,626	1,614	1,613	1,595	1,617	1,535
R-sq	0.331	0.319	0.302	0.334	0.223	0.312	0.245	0.323
Nurse–Shift FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hospital–Hour FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ambulance FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Chief Complaint FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Visit Number FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Outcomes capture diagnostic information collection during the emergency department visit after triage: any imaging ordered and number of distinct imaging tests (Cols. 1–2), any laboratory tests and number of distinct lab tests (Cols. 3–4), any vitals reassessed and number of vitals reassessments (Cols. 5–6), and any pain reassessment and number of pain reassessments (Cols. 7–8). The omitted group (non-White patients under low time pressure) has an average outcome of 69.40% in Col. 1, 2.20 in Col. 2, 87.45% in Col. 3, 43.16 in Col. 4, 80.44% in Col. 5, 6.72 in Col. 6, 49.53% in Col. 7, and 6.78 in Col. 8. "High Time Pressure (High TP)" equals one if the triage nurse was in the top two bins of any of the three time pressure proxies (number of patients waiting, next-patient wait time, or previous triage duration; see text). Controls include a male indicator and a cubic in age. All models include nurse×shift, hospital-hour, arrival mode, chief complaint, and visit number fixed effects. Standard errors (in parentheses) are clustered by triage nurse. Sample includes all ESI 1–3 patients.

Table A.4: Time Pressure and Racial Disparities in Downstream Treatment

	Procedures		Pain Meds		Hosp. Admission
	(1)	(2)	(3)	(4)	(5)
	Any	Number	Any	Number	Any
High Time Pressure (TP)	-0.005***	-0.007	-0.005***	-0.022	-0.007***
	(0.001)	(0.005)	(0.001)	(0.017)	(0.001)
Non-Hispanic White (Low TP)	0.036***	0.094***	0.005***	0.410***	0.046***
	(0.002)	(0.007)	(0.002)	(0.024)	(0.003)
High TP × White	0.006***	0.018**	0.005***	0.051**	0.016***
	(0.002)	(0.007)	(0.002)	(0.026)	(0.002)
Observations	1,278,090	638,330	1,265,674	406,582	1,093,712
Nurses	1,613	1,549	1,617	1,431	1,610
R-sq	0.354	0.249	0.261	0.316	0.359
Nurse–Shift FE	Yes	Yes	Yes	Yes	Yes
Hospital–Hour FE	Yes	Yes	Yes	Yes	Yes
Ambulance FE	Yes	Yes	Yes	Yes	Yes
Chief Complaint FE	Yes	Yes	Yes	Yes	Yes
Visit Number FE	Yes	Yes	Yes	Yes	Yes

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Outcomes capture treatment decisions during the emergency department visit after triage: any procedures and number of distinct procedures (Cols. 1–2), any pain medication and number of pain medications (Cols. 3–4), and hospital admission (Col.5). The omitted group (non-White patients under low time pressure) has an average outcome of 55.33% in Col. 1, 1.74 in Col. 2, 36.85% in Col. 3, 3.22 in Col. 4, and 23.32% in Col. 5. "High Time Pressure (High TP)" equals one if the triage nurse was in the top two bins of any of the three time pressure proxies (number of patients waiting, next-patient wait time, or previous triage duration; see text). Controls include a male indicator and a cubic in age. All models include nurse×shift, hospital-hour, arrival mode, chief complaint, and visit number fixed effects. Standard errors (in parentheses) are clustered by triage nurse. Sample includes all ESI 1–3 patients.