Appendix

A. ED Arrival Patterns and Capacity

Figure A1: Arrival Rate by Hour

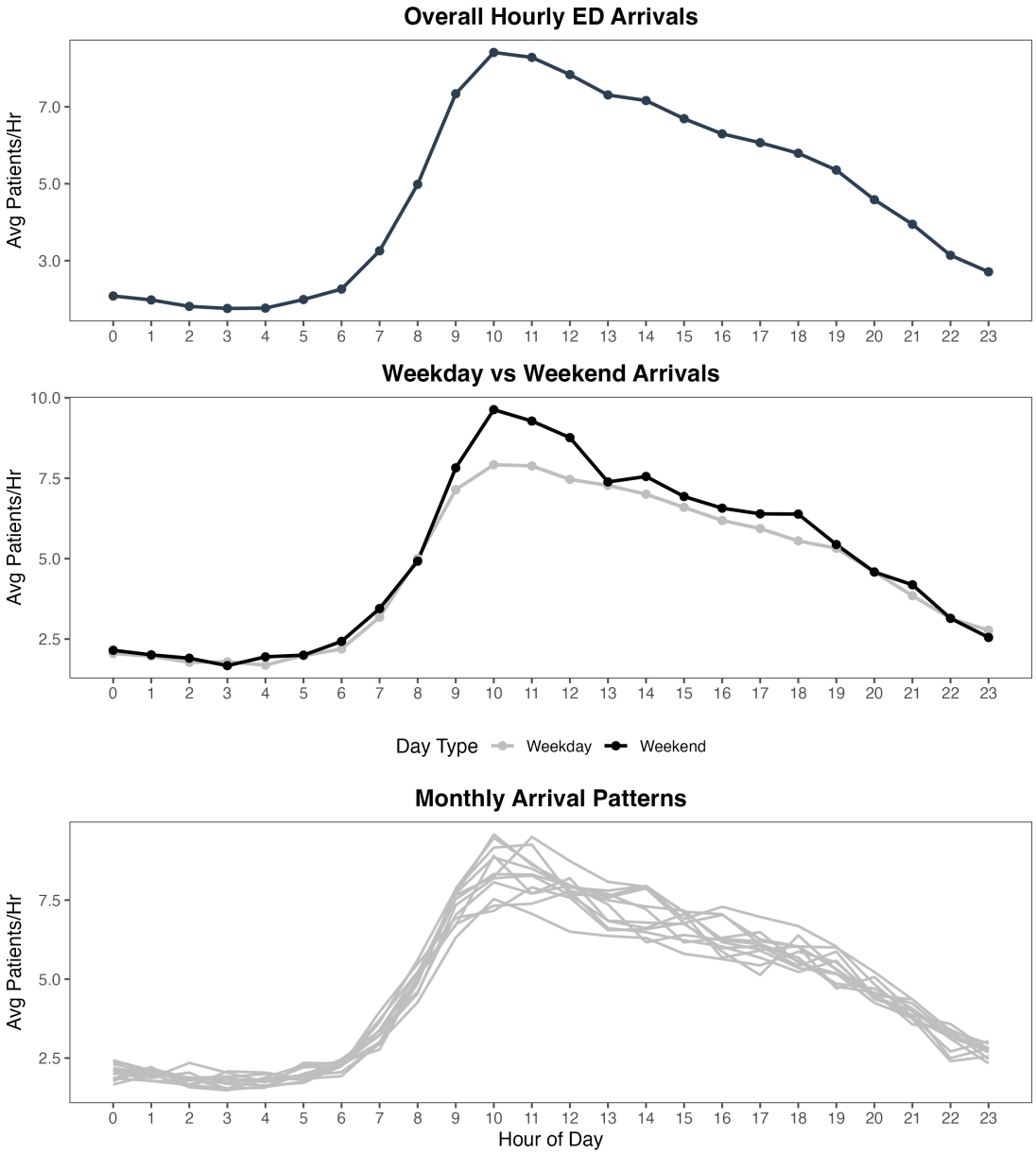
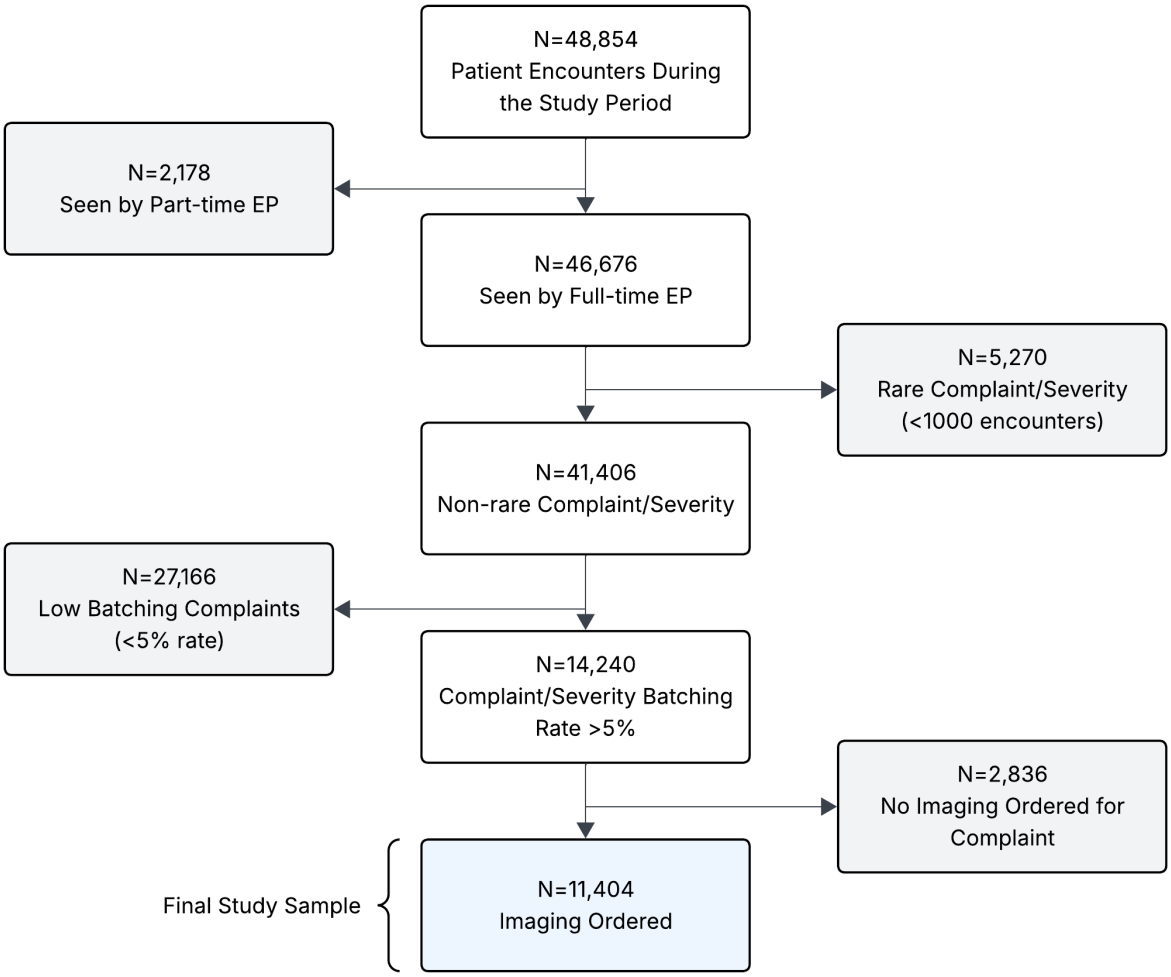


Table A1: Internal Guidelines for Activation of Overcapacity and Saturation Plans for the Emergency Department (ED)

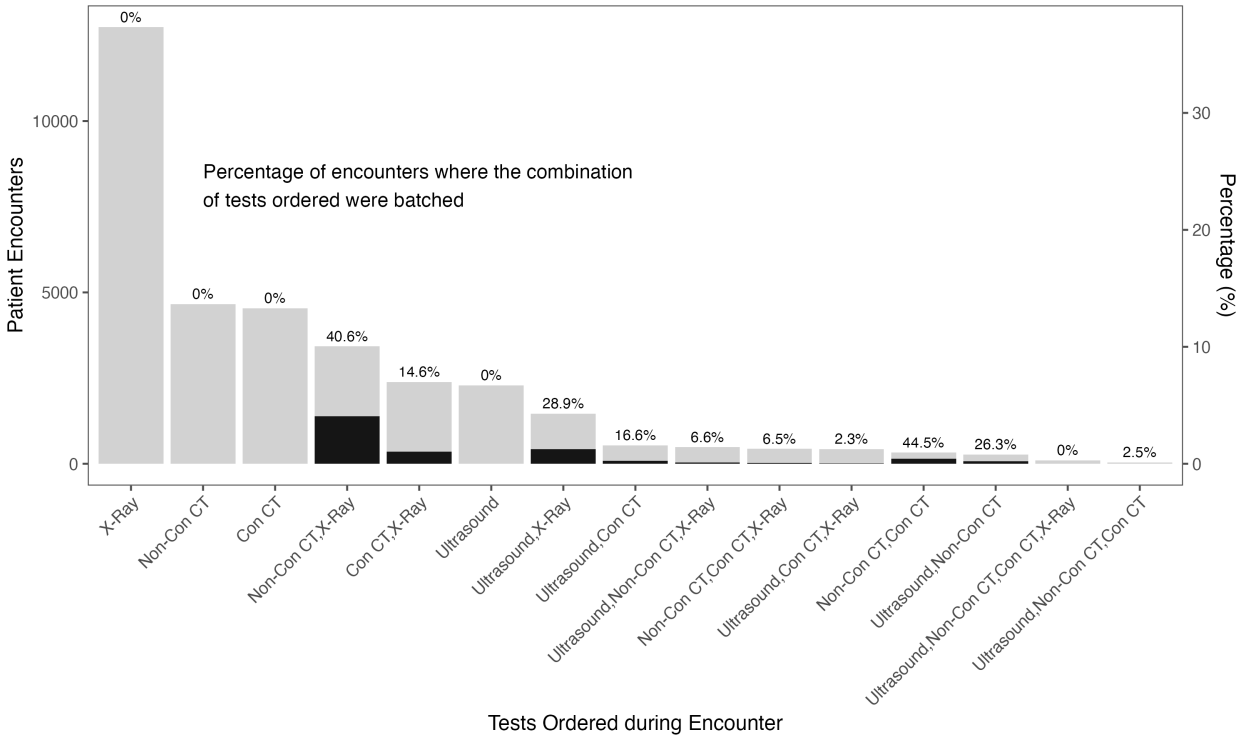
Capacity	Conditions	Actions
Normal Operations	1. Patients in the waiting room for less than 20 minutes 2. Number of patients in the ED is less than available ED beds	1. No utilization of Rapid Medical Assessment (RMA) resources or waiting room evaluations unless the patient will not require a treatment room. 2. No laboratory tests to be done from the waiting room as patients are being roomed immediately. 3. Nurse-initiated protocols to be entered by the throughput nurse (TN) or bedside RN 30 minutes after patient arrival, regardless of patient location.
Minor Overcapacity	1. Waiting room time 21–90 minutes OR 2. Waiting room with 10 patients OR 3. Number of ED patients exceeds available ED beds by 10 (e.g., admitted patients waiting for a bed assignment) OR 4. Team Lead (TL) discretion	1. All actions for Normal Operations plus: 2. Laboratory testing to be initiated in the ED waiting room by ED staff. 3. RMA workflow activation with dedicated staff for management of waiting room patients. 4. All zones available for placement of patients for all providers. 5. TN communicates with consulting providers to expedite care.
Major Overcapacity	1. Waiting room time greater than 90 minutes OR 2. Waiting room with more than 20 patients OR 3. More than 40 new patient arrivals in 2 hours OR 4. Number of ED patients exceeds available ED beds by 20 OR 5. TL discretion	1. All actions for Normal Operations and Minor Overcapacity plus: 2. TL to consider placing ED on diversion. 3. Activation of RMA workflow with a second RMA if capacity allows. 4. External/hospital-wide levers activated (e.g., pop-ups by house supervisor, transport crisis, expedited bed assignments, priority placement of ED patients). 5. Activation of Obs Unit APPs to ED if observation duties allow. 6. Activation of on-call ED Physician if physician cap criteria is met.

Figure A2: CONSORT Flow Diagram for Sample Selection



Notes: This figure displays the sample selection process for the Mayo Clinic ED data. Starting with 48,854 patient encounters during the study period (October 2018 - December 2019), we apply sequential exclusion criteria to arrive at our analytical sample of 11,404 encounters. Exclusions are necessary to ensure sufficient variation in batching behavior for instrumental variable identification. Rare complaints are those with fewer than 1,000 total encounters. Low-batching complaints are those where batching occurs in less than 5% of encounters.

Figure A3: Distribution of imaging test combinations and batching rates



Notes:

B. Complier, Always-Taker, and Never-Taker Classification

In this section, we describe the method used to calculate the share of compliers (as well as always-takers and never-takers) in the context of batch ordering in the emergency department (ED). The method follows [Dahl et al., 2014, Dobbie et al. [2018], and Eichmeyer and Zhang [2022]].

First, compliers are defined as patients who would not have had their imaging tests ordered in a batch if they had been seen by a sequencer (low-batching physician) but would have had their imaging tests batched if they had been seen by a batcher (high-batching physician):

$$\pi_{\text{complier}} = P(B_{\bar{z}i} > B_{\underline{z}i}) = \mathbb{E}(B_{\bar{z}i} - B_{\underline{z}i}) = P(B_i|Z_i = \bar{z}) - P(B_i|Z_i = \underline{z})$$

where B_i represents the batch ordering decision for patient i, Z_i represents the batch tendency of patient i's physician, and $\bar{z}$ and z represent the highest and lowest batch tendency physicians, respectively.

Similarly, always-takers are patients whose imaging tests would be batched regardless of which physician they see:

$$\pi_{\text{always-taker}} = P(B_{\bar{z}i} = B_{\underline{z}i} = 1) = P(B_{\underline{z}i} = 1)$$

where the last step follows from the monotonicity assumption. Conversely, never-takers are patients whose imaging tests would never be batched regardless of which physician they see:

$$\pi_{\text{never-taker}} = P(B_{\bar{z}i} = 0)$$

To estimate these proportions, we define the most aggressive batch-ordering physicians ($\bar{z}$) as those in the top percentile of batch tendency and the most conservative batch-ordering physicians ($\underline{z}$) as those in the bottom percentile. We then compute the share of compliers, always-takers, and never-takers using first-stage moments. Specifically, we fit a local linear regression of $Batched_i$ on physician batch tendency, take the share of patients who receive batched imaging under the top percentile of batch tendency, and subtract the share of patients who receive batched imaging under the bottom percentile.

Additionally, we can characterize our compliers based on observable characteristics. For example, we can calculate the share of patients with specific clinical or demographic features who are compliers. In particular, we can compute:

$$P(X_i = x|\text{complier})$$

This framework allows us to understand which patients are most likely to be affected by variation in physician batching behavior and how batch ordering decisions propagate through ED operational processes.

$$P(X_i = x|\text{complier}) = P(X_i = x|B_{\bar{z}i} > B_{\underline{z}i})$$

$$= \frac{P(X_i = x \cap B_{\bar{z}i} > B_{zi})}{P(B_i|Z_i = \bar{z}) - P(B_i|Z_i = \underline{z})}$$

$$= \frac{P(B_{\bar{z}i} > B_{\underline{z}i}|X_i = x)P(X_i = x)}{\pi_{\text{complier}}}$$

$$= \frac{\pi_{c|x}P(X_i = x)}{\pi_c}$$

This moment is calculated by computing the share of compliers for the subsample where $X_i = x$ (i.e., checking the moments of the first stage for that subsample) and scaling it by the unconditional share of that subsample, divided by the overall share of compliers.

C. Mediation Analysis: SEM Results

Table C1: Mediation Analysis of the Effect of Batching on ED Length of Stay (LOS) via Imaging and Admission Decisions

	Estimate	Standard Error	p-value
Number of Imaging Tests (Residualized)			
Batching (b_1)	1.434	0.106	<0.001
Tachycardic	0.077	0.015	<0.001
Tachypneic	0.111	0.020	<0.001
Febrile	0.051	0.028	0.062
Hypotensive	0.111	0.038	0.003
Arrival Age (scaled)	0.002	0.000	<0.001
Admission Decisions (Residualized)			
Batching (b_2)	0.327	0.069	<0.001
Number of Imaging Tests (b_3)	0.162	0.006	<0.001
Tachycardic	0.084	0.010	<0.001
Tachypneic	0.056	0.013	<0.001
Febrile	0.142	0.018	<0.001
Hypotensive	0.210	0.024	<0.001
Arrival Age (scaled)	0.002	0.000	<0.001
Length of Stay (Residualized)			
Number of Imaging Tests (c_1)	0.144	0.006	<0.001
Admission Decisions (c_2)	0.089	0.010	<0.001
Batching Direct Effect (c')	0.031	0.072	0.662
Tachycardic	0.041	0.010	<0.001
Tachypneic	-0.005	0.013	0.721
Febrile	-0.020	0.018	0.283
Hypotensive	-0.025	0.025	0.326
Arrival Age (scaled)	0.002	0.000	<0.001
Indirect and Total Effects			
Indirect via Imaging Tests ($b_1 \times c_1$)	0.207	0.018	<0.001
Indirect via Admission Decisions ($b_2 \times c_2 + b_1 \times b_3 \times c_2$)	0.050	0.008	<0.001
Total Indirect Effect	0.256	0.020	<0.001
Total Effect	0.288	0.073	<0.001

Notes: This table presents the results of a structural equation model (SEM) investigating the relationships between batching, the number of imaging tests, admission decisions, and ED length of stay (LOS). To address potential confounding from fixed effects (e.g., complaint type and time of month), all variables were residualized by regressing them on the fixed effects prior to the SEM. This residualization ensures that the estimates reflect associations net of the fixed effects.

This analysis is exploratory, aiming to understand the mechanism by which batching influences ED operational outcomes. While the directions and significance of these associations align with hypothesized pathways, these results should not be interpreted as causal mediation but rather as an exploratory investigation of plausible mechanisms. The findings suggest that increased diagnostic intensity is a key pathway by which batching extends LOS, with admissions playing a secondary role.

Table C2: Mediation Analysis of the Effect of Batching on Time to Disposition

	Estimate	Standard Error	p-value
Number of Imaging Tests (Residualized)			
Batching (b_1)	1.434	0.106	<0.001
Tachycardic	0.077	0.015	<0.001
Tachypneic	0.111	0.020	<0.001
Febrile	0.051	0.028	0.062
Hypotensive	0.111	0.038	0.003
Arrival Age (scaled)	0.002	0.000	<0.001
Time to Disposition (Residualized)			
Number of Imaging Tests (c_1)	0.060	0.007	<0.001
Batching Direct Effect (c')	-0.100	0.081	0.216
Tachycardic	0.006	0.011	0.617
Tachypneic	-0.018	0.015	0.236
Febrile	-0.098	0.021	<0.001
Hypotensive	-0.090	0.028	0.002
Arrival Age (scaled)	-0.001	0.000	0.042
Indirect and Total Effects			
Indirect via Imaging Tests ($b_1 \times c_1$)	0.085	0.012	<0.001
Total Effect	-0.015	0.081	0.855

Notes: This table presents results from a structural equation model (SEM) analyzing the relationship between batching, the number of imaging tests, and time to disposition. All variables were residualized to account for fixed effects of day of the week, month, and chief complaint by severity. The coefficients for the number of imaging tests reflect its role as a mediator in the pathway from batching to time to disposition. The direct effect of batching is not significant, while the indirect effect via imaging tests is significant and positive, supporting the hypothesis that increased diagnostic intensity prolongs time to disposition. This analysis is exploratory and intended to provide insight into plausible mechanisms, not causal mediation.

D. Robustness Checks

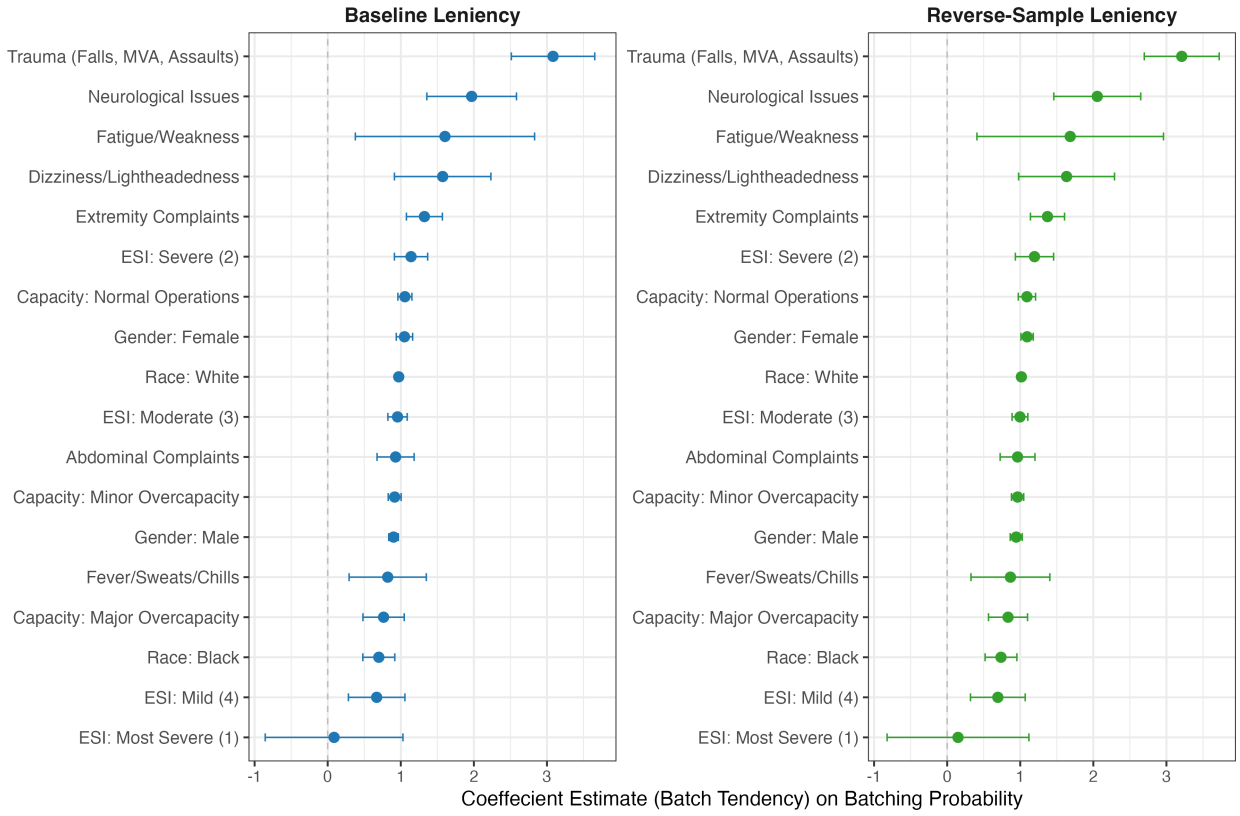


Figure D1: Testing the Monotonicity Assumption

Notes: The left panel displays the first stage coefficient and 95% CI of batching on the baseline physician batch tendency instrument for the corresponding sub-sample. The right panel constructs a new physician batch tendency instrument using all emergency visits, excluding the corresponding sub-sample ("reverse-sample"), and displays the coefficient and 95% CI of the first stage regression back on that sub-sample. Robust standard errors are clustered at the physician level.

Table D1: Placebo Check: Batch Tendency and Patient Outcomes

	Dependent variable		
	Log time to disposition (1)	Log LOS (2)	72hr return with admission (3)
Batch tendency	3.486	3.580	−0.2158
	(1.724)	(1.750)	(0.1069)
Mean dependent variable	4.57	4.59	0.001
	(0.611)	(0.615)	(0.098)
Time FE	Yes	Yes	Yes
Baseline controls	Yes	Yes	Yes
Adj R^2	0.159	0.160	0.051
Observations	1,244	1,244	1,244

Notes: This table reports the estimated coefficients of a reduced-form regression of our main outcomes on physician batch tendency for patients with conditions for which batching occurs less than 1% of the time. The dependent variables are log time to disposition, log LOS, and 72-hour return with admission. Robust standard errors are clustered at the physician level.

Placebo Exercise

We investigate whether the reduced-form effects observed in Section 4.1 are due to differences in batch ordering rates across physicians or due to other provider differences correlated with batching tendency. We examine this by studying reduced-form effects among patients with conditions that rarely require multiple imaging tests, as a falsification check. For example, consider a patient who arrives at the ED with an isolated ankle injury—a condition for which multiple imaging tests are rarely ordered. For such patients, we should expect to see no impact of physician batch tendency if "batchers" and "sequencers" do not systematically differ in other dimensions of care relevant to patient outcomes.

We restrict attention to ED visits for conditions which batching occurs less than 1 percent of the time, as a placebo check. We estimate reduced-form regressions of each main outcome on physician batch tendency for this subsample and the results are presented in Table D1. The results show no significant association between physician batch tendency and patient outcomes for this subsample, providing evidence that the reduced-form effects observed in the main analysis are not driven by unobserved physician differences correlated with batching tendency.

This placebo analysis supports our identification strategy by suggesting that batch tendency is not systematically correlated with other physician practice patterns that might affect patient outcomes. However, we recognize that given the multidimensionality of physician behavior, we also present reduced-form results and maintain focus on imaging-related outcomes, for which such concerns are less pronounced.

Table D2: Robustness Check for Main 2SLS Results

	(1)	(2)
Panel A. Primary Outcomes		
Log time to disposition	0.923^{*}	$0.753^{\dagger}$
	(0.358)	(0.368)
Log LOS	0.783^{*}	0.656^{*}
	(0.300)	(0.296)
Number of distinct imaging tests	1.271^{***}	1.307^{***}
	(0.218)	(0.213)
72hr return with admission	-0.005	-0.015
	(0.018)	(0.021)
Time FE	Yes	Yes
Baseline controls	Yes	Yes
Admit	Yes	No
Testing tendency	No	Yes
Observations	11,404	11,404

Notes: This table reports the 2SLS estimates of the effect of batching on patient outcomes, using specifications that include the additional controls of admitting status and testing tendency. $^{\dagger}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

References

Gordon B. Dahl, Andreas Ravndal Kostøl, and Magne Mogstad. Family welfare cultures. *The Quarterly Journal of Economics*, 129(4):1711–1752, November 2014. doi: 10.1093/qje/qju019. URL https://doi.org/10.1093/qje/qju019.

Will Dobbie, Jacob Goldin, and Crystal S. Yang. The effects of pretrial detention on conviction, future crime, and employment: Evidence from randomly assigned judges. *American Economic Review*, 108(2):201–240, 2018.

Sarah Eichmeyer and Jonathan Zhang. Pathways into opioid dependence: Evidence from practice variation in emergency departments. *American Economic Journal: Applied Economics*, 14(4):271–300, 2022. doi: 10.1257/app.20210048.