

INFORMS 2017

Michael Freeman
INSEAD

Gatekeeping under Congestion: An Empirical Study of Referral Errors in the Emergency Department

Joint work with:

Susan Robinson – Cambridge University Hospitals

Stefan Scholtes – Cambridge Judge Business School

Decision making in the ED

- Emergency providers make disposition decisions ~350,000 times/day in US EDs
 - Option 1: admit patient to hospital
 - Option 2: discharge patient home
- ED physicians act as **gatekeepers** to inpatient beds
- Emergency medicine: High levels of clinical uncertainty and variation in diagnostic accuracy
- ED physicians under time and workload induced pressure

Research questions

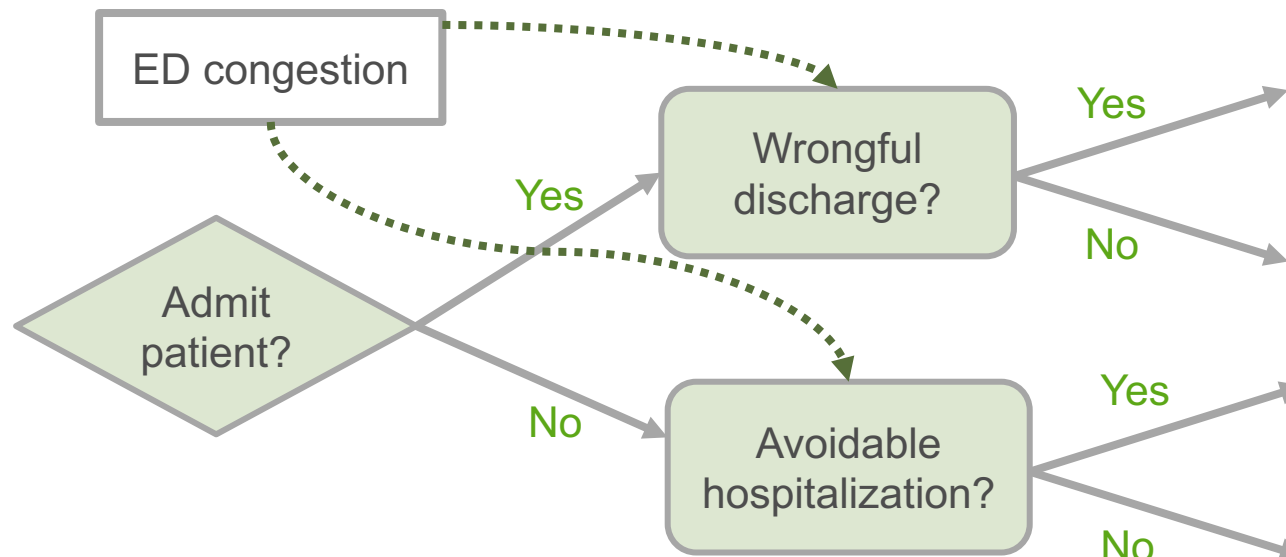
RQ1: How does congestion affect the accuracy of ED gatekeeping decisions?

- **Wrongful discharges** (false negative)
- **Avoidable hospitalizations** (false positive)

Research questions

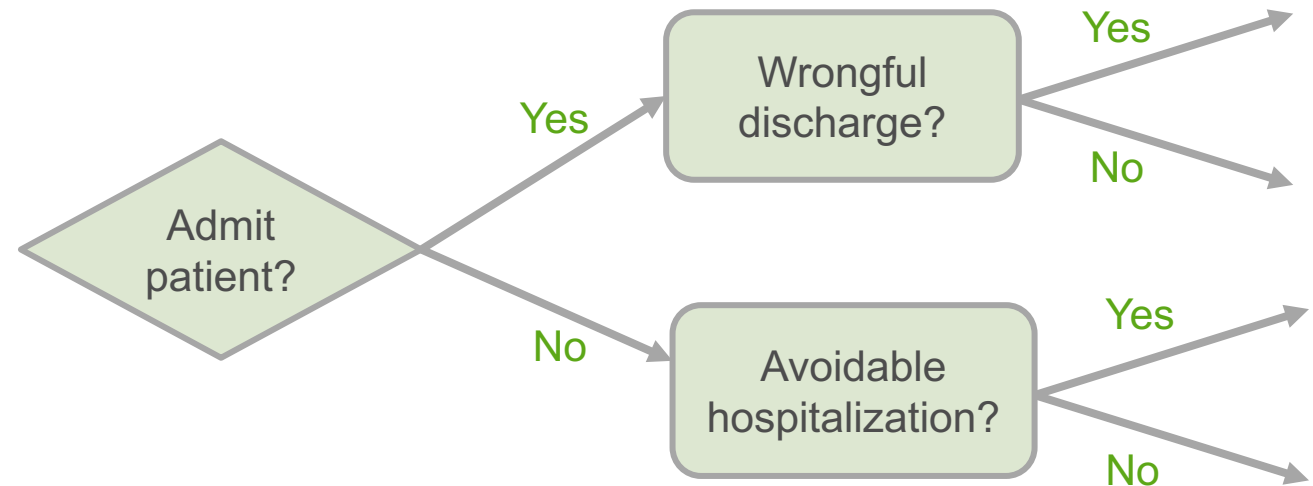
RQ1: How does congestion affect the accuracy of ED gatekeeping decisions?

- **Wrongful discharges** (false negative)
- **Avoidable hospitalizations** (false positive)



Research questions

RQ2: Are there process changes that can reduce the rate of false negatives and false positives?

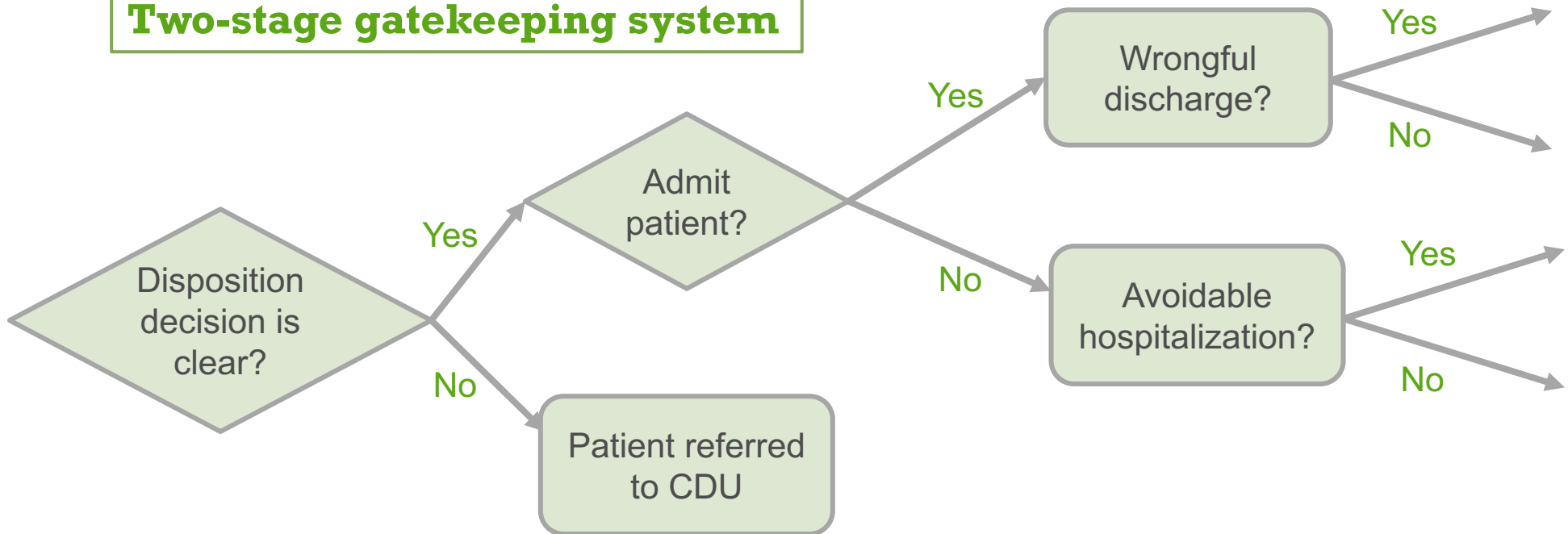


Single-stage gatekeeping system

Research questions

RQ2: Are there process changes that can reduce the rate of false negatives and false positives?

Two-stage gatekeeping system



- **Gatekeeping**

- Shumsky & Pinker (2003) and Hasija et al. (2005): contracting for system optimal rate of referrals
- Lee et al. (2012) and Zhang et al. (2011): outsourcing contracts and security-check performance
- Freeman et al. (2016): study of workload induced changes in gatekeeping decisions in maternity care

- **Type classification**

- Argon & Ziya (2009): model customer classification policies with imperfect information about customer type
- Alizamir et al. (2013) and Wang et al. (2010): increasing service duration can improve accuracy of diagnosis, at cost of increased congestion

- **Speed, quality & load**

- Anand et al. (2011) and Kostami & Rajagopalan (2013): service value increasing in duration, but cost of wait
- Hopp et al. (2007): increasing capacity may increase congestion as discretionary service components added
- Debo et al. (2008) and Paç & Veeraraghavan (2015): congestion acts as a deterrent to expert overtreatment

- **Streaming/triage**

- Saghafian et al. (2012, 2014, 2017): triage can be augmented to stream ED patients based not only on their severity but also e.g. their complexity

- All visits to the ED of a large UK-based teaching hospital over a 7 year period (2007-2013)
 - Approx. 250 patient visits per day; 29% of patients admitted
 - Data set includes all inpatient info associated with those patients admitted via the ED
 - Subset to approx. 500k obs. of patients over the age of 16 who did not leave without being seen
- Exclude dates around public holidays when ED staffing and patient arrivals atypical
- Use first year of data as a run-in period to generate controls and measures of patient risk
- Leaves ~**374k** obs. for analysis

Dependent variables

(RQ1: How does congestion affect the accuracy of ED gatekeeping decisions?)

- **Wrongful discharges**

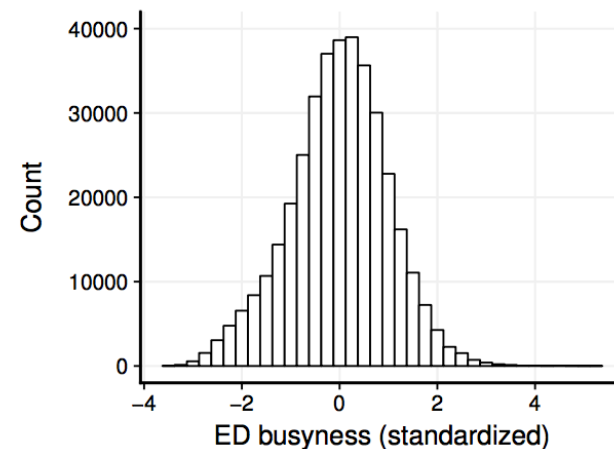
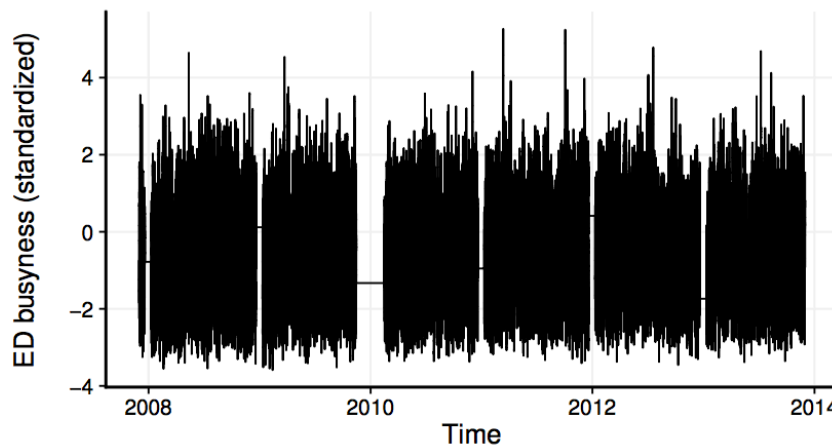
- Patient discharged home from ED but revisits ED within 7 days and is then admitted to the hospital
- 1.0% of ED visits and 1.5% of all patients discharged

- **Avoidable hospitalizations**

- Patient admitted to an inpatient unit and discharged within 24hrs without treatment
- 4.3% of ED visits and 13.7% of all admissions
- Change in rate indicative of change in likelihood of false admissions

Measuring congestion

- Observed occupancy in the ED at hour h : $QueueED_h$
- Proxy capacity by predicting using quantile regression the expected 95th percentile of occupancy: $QueueED_h^{95th}$
 - In QR, control for temporal effects (year, quarter, weekend, hour in 4-hour windows)
- ED congestion = $QueueED_h / QueueED_h^{95th}$
- Standardize by subtracting the mean and dividing by the SD



Hypothesis 1

- ED physicians are congestion-sensitive gatekeepers who:
 - are aware of congestion levels in the GK system
 - adjust their service to trade off time spent with an individual patient versus improved system throughput
- Time pressure reduces accuracy of diagnosis (e.g. Alizamir et al. 2013) meaning that ED physicians have to make gatekeeping decisions under increased clinical uncertainty

Hypothesis 1: As system congestion increases, congestion-sensitive gatekeepers make more errors in their referral decisions

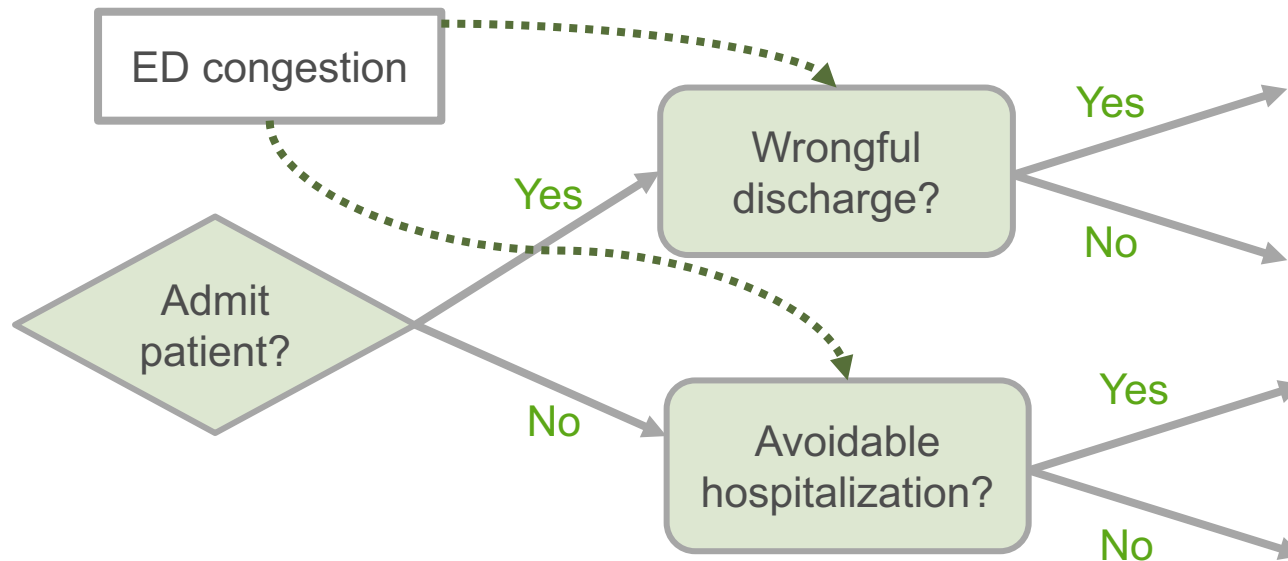
Hypothesis 2

- Gatekeepers who have asymmetric disutilities for the two types of errors can trade them off against one another

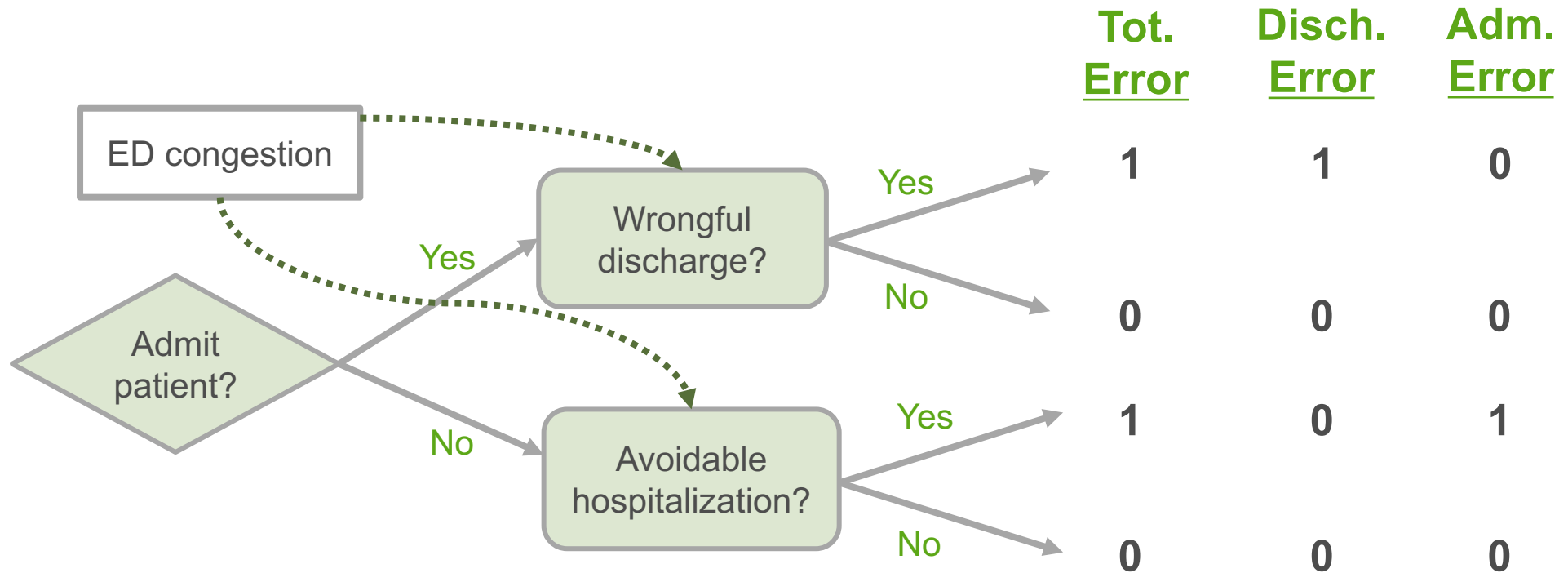
Hypothesis 2: If congestion-sensitive gatekeepers weigh one type of referral error more heavily than the other, the proportion of the more heavily weighted error will fall with system congestion

- Expect ED physicians to weigh the cost of a wrongful discharge more heavily than an avoidable hospitalization: predict an increase in avoidable hospitalizations relative to wrongful discharges

Model: Standard probit



Model: Standard probit

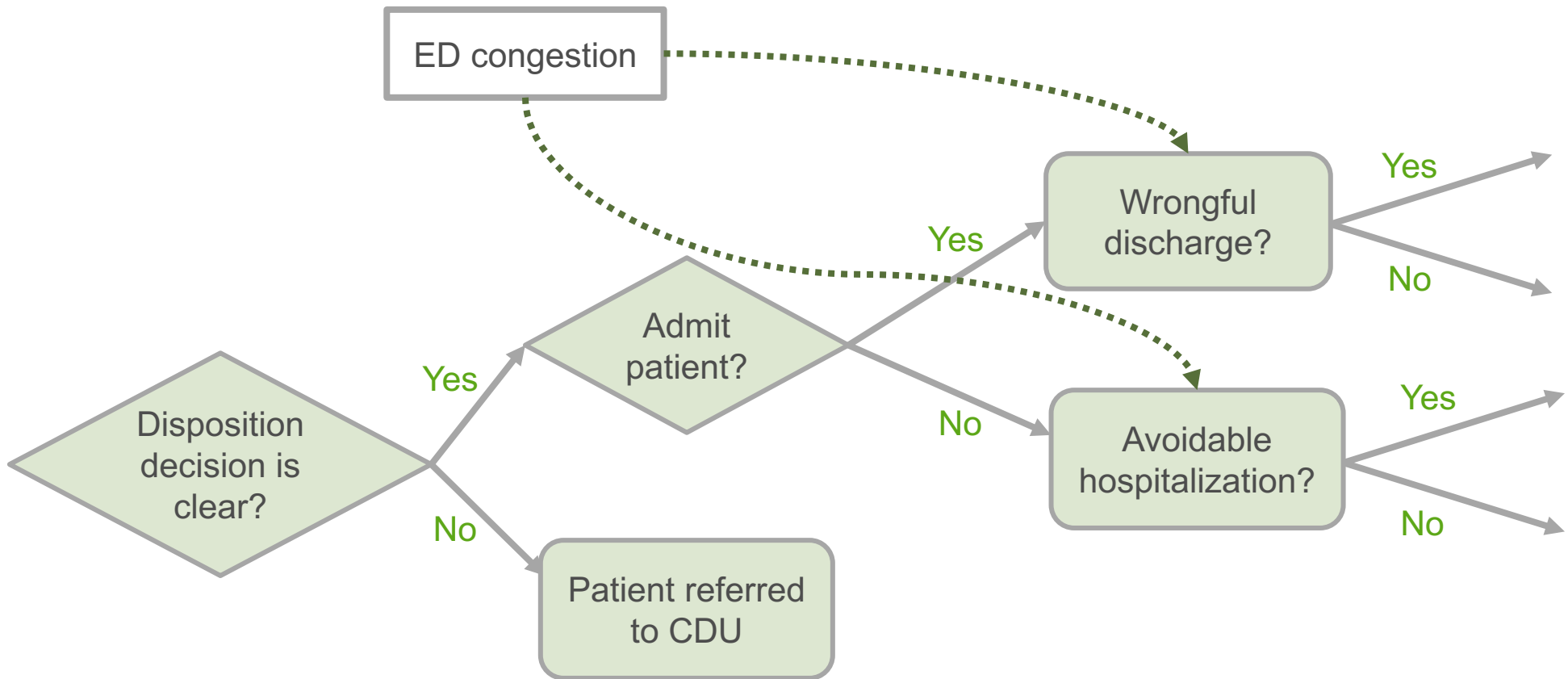


$$y_i^* = \alpha + u_i\beta + x_i\gamma + \varepsilon_i, \quad \varepsilon_i \sim N(0,1)$$

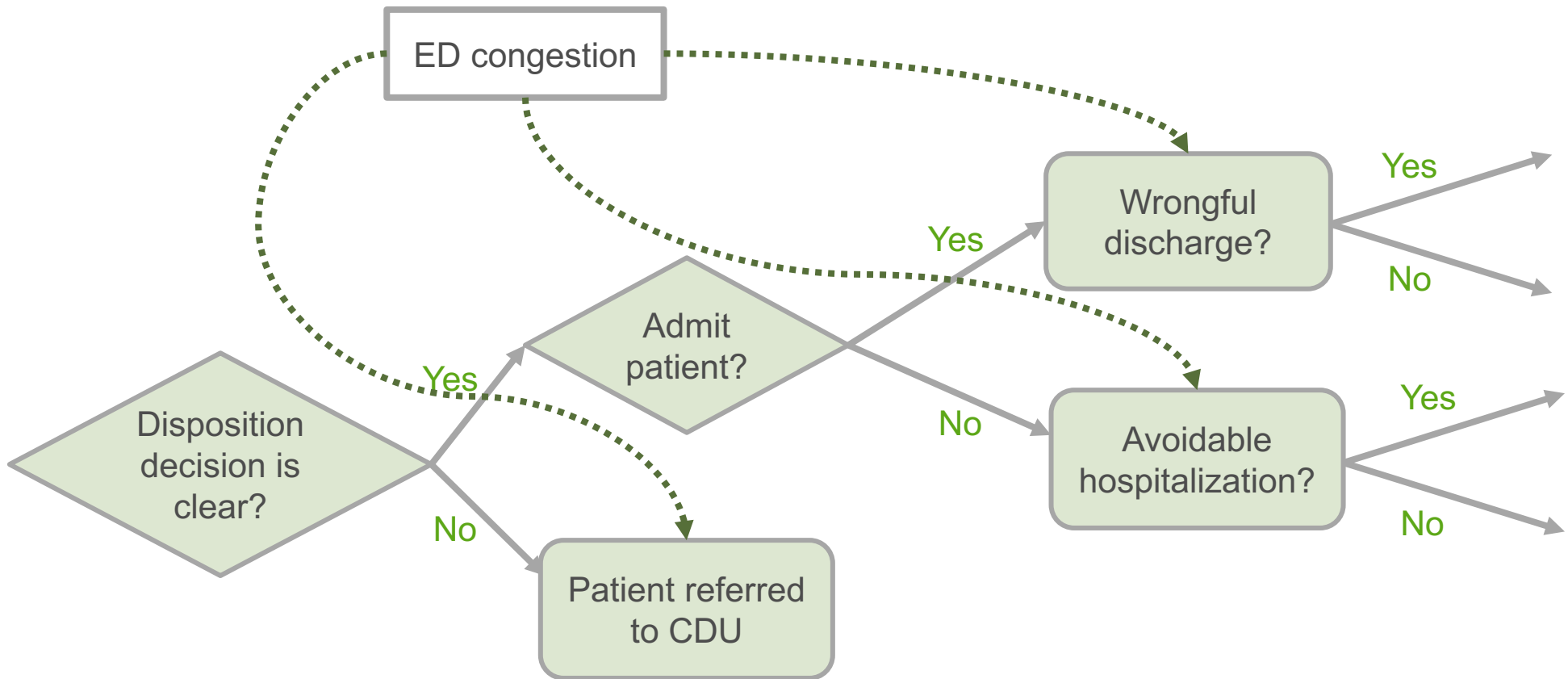
$$y_i^{\text{outcome}} = 1(y_i^* > 0)$$

(where u_i is ED crowding)

Modeling challenge



Modeling challenge



Heckman probit model

$$\begin{aligned} y_i^{\text{outcome}} &= 1(\alpha + u_i\beta + x_i\gamma + \varepsilon_i > 0) \\ y_i^{\text{select}} &= 1(\alpha' + u_i\beta' + x_i\gamma' + z_i\delta' + \varepsilon_i' > 0) \\ (\varepsilon_i, \varepsilon_i') &\sim N \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}. \end{aligned}$$

(where u_i is ED crowding)

- $y_i^{\text{select}} = 1$ if the patient is **not** referred to the CDU
- Censoring assumption: y_i^{outcome} is not observed when $y_i^{\text{select}} = 0$
- Consistent and asymptotically efficient coefficient estimates when $\rho \neq 0$ (Van de Ven / Van Praag 1981)

Model estimation improved and coefficients more reliable when IVs provided

IV 1: Propensity of assigned ED physician to refer patients into CDU over previous year

Relevance:

- If a patient is assigned to a physician who has a history of referring more patients to the CDU, they are more likely to be referred there themselves

Validity:

- Include controls for the physician's historic error rates over the same period
- After controlling for physician-specific historic error rates, the CDU referral rate of the assigned physician is uncorrelated with the residuals

Model estimation improved and coefficients more reliable when IVs provided

IV 2: Congestion level in the CDU

Relevance:

- If the CDU is congested then it becomes less available to ED physicians as an option

Validity:

- For patients who are not admitted to the CDU, the busyness of the CDU should have no direct effect on their likelihood of being admitted or discharged in error

(Instrument validity can be tested with several instruments → all tests pass)

Congestion effect

	Decision made by ED physicians		
	(1e) TotErr	(2e) ObsAdm	(3e) DischErr
ED congestion	0.020*** (0.005)	0.028*** (0.005)	−0.014 [†] (0.008)
ρ	−0.220*** (0.040)	−0.114* (0.047)	−0.179** (0.057)
N	373,663	373,663	373,663
N uncensored	337,144	337,144	337,144
Log-lik	−157,985	−147,023	−116,471

- A 1 σ increase in ED congestion results in a:
 - 7.2% increase in both types of error
 - 7.7% increase in avoidable hospitalizations
 - 3.3% reduction in wrongful discharges

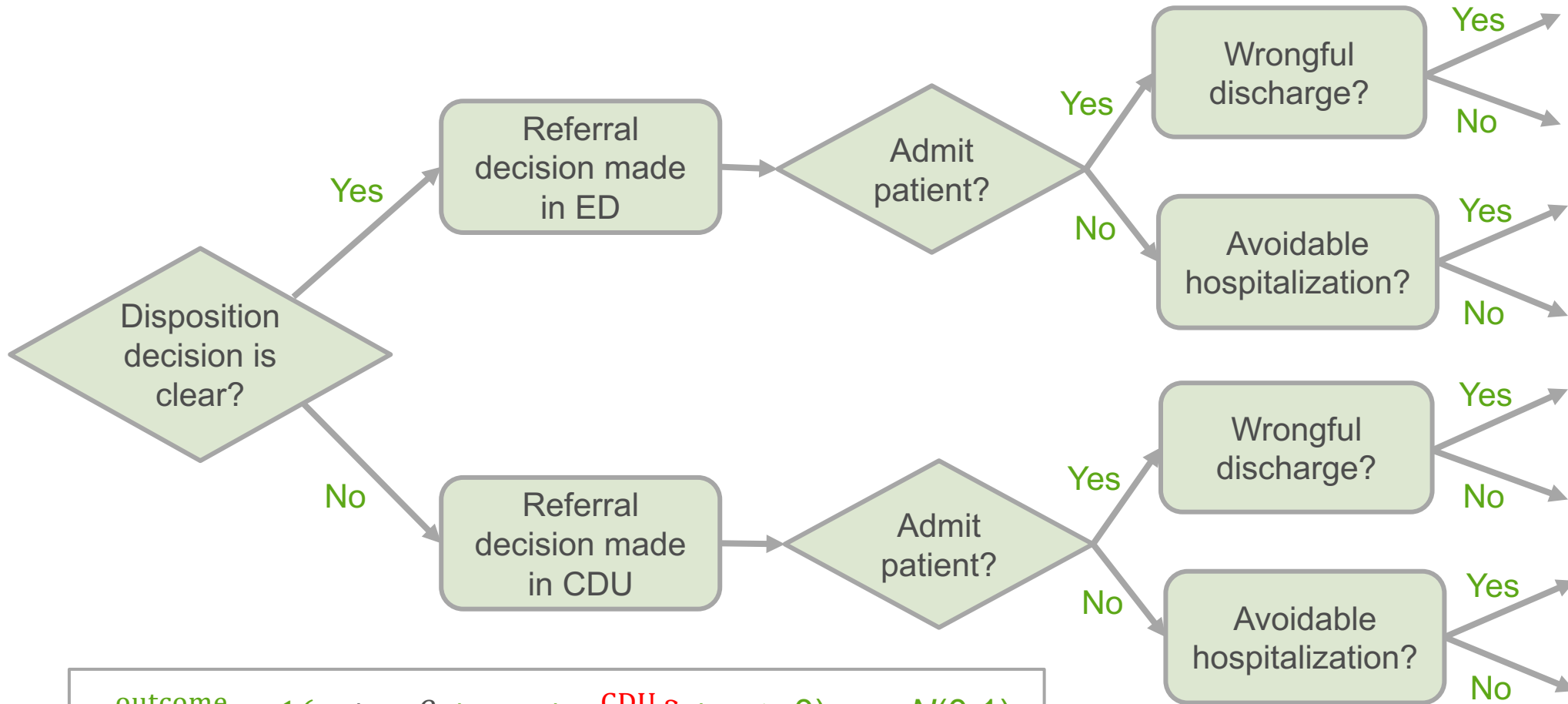
Consistent with
Hypotheses 1
and 2

Hypothesis 3

- The two-stage GK system improves the match between customers with heterogeneous needs with gatekeepers with heterogeneous experience and resources

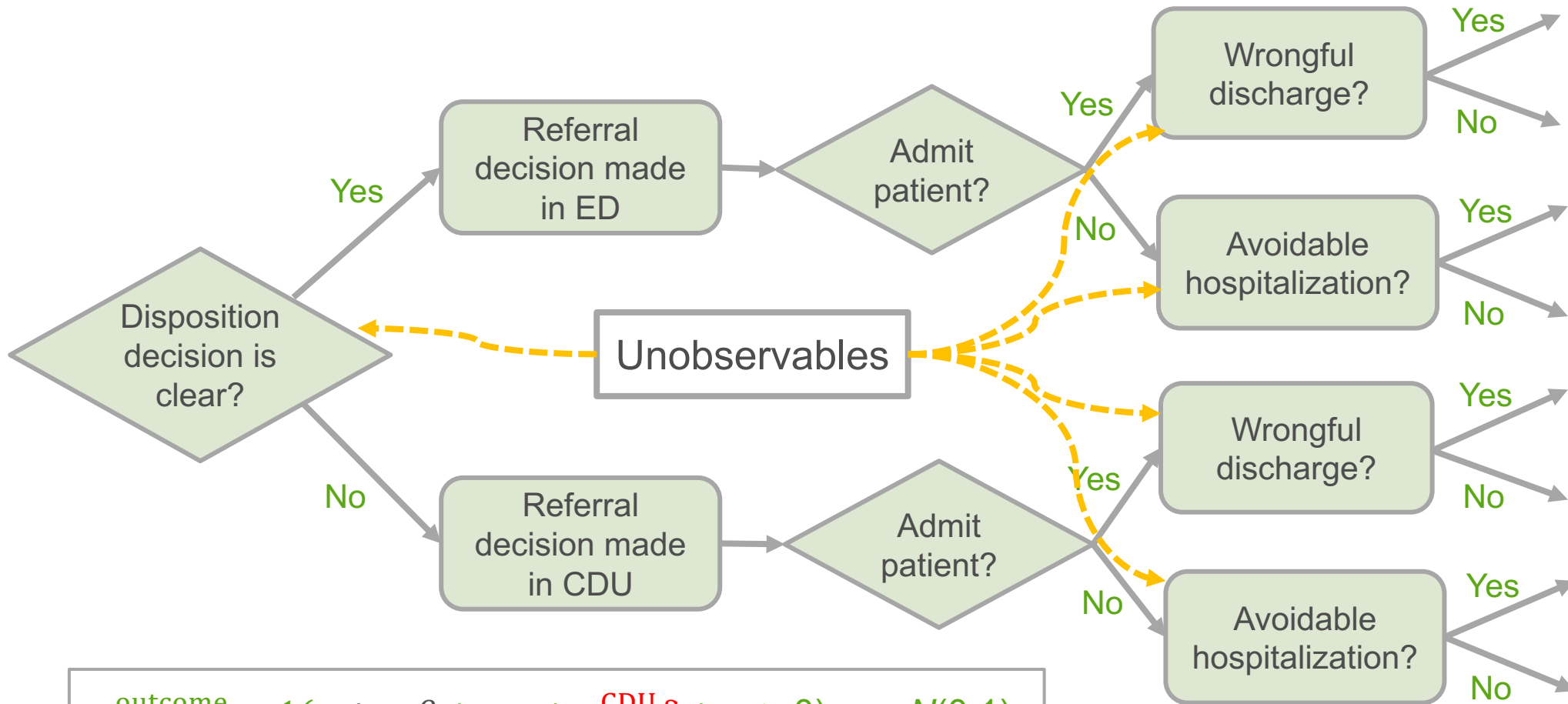
Hypothesis 3: A two stage gatekeeping system reduces both types of gatekeeping error

Two-stage gatekeeping system



$$y_i^{\text{outcome}} = 1(\alpha + u_i\beta + x_i\gamma + y_i^{\text{CDU}}\delta + \varepsilon_i > 0), \quad \varepsilon_i \sim N(0,1)$$

Two-stage gatekeeping system



$$y_i^{\text{outcome}} = 1(\alpha + u_i\beta + x_i\gamma + y_i^{\text{CDU}}\delta + \varepsilon_i > 0), \quad \varepsilon_i \sim N(0,1)$$

Recursive bivariate probit

Heckprobit

$$\begin{aligned}y_i^{\text{outcome}} &= 1(\alpha + u_i\beta + x_i\gamma + \varepsilon_i > 0) \\y_i^{\text{select}} &= 1(\alpha' + u_i\beta' + x_i\gamma' + z_i\delta' + \varepsilon'_i > 0) \\(\varepsilon_i, \varepsilon'_i) &\sim N\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.\end{aligned}$$

Effect of interest: impact of congestion on referral decisions of the first-stage GK

Coefficient of interest: β

Recursive bivariate probit

$$\begin{aligned}y_i^{\text{outcome}} &= 1(\alpha + u_i\beta + x_i\gamma + y_i^{\text{CDU}}\delta + \varepsilon_i > 0) \\y_i^{\text{CDU}} &= 1(\alpha' + u_i\beta' + x_i\gamma' + z_i\delta' + \varepsilon'_i > 0) \\(\varepsilon_i, \varepsilon'_i) &\sim N\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.\end{aligned}$$

Effect of interest: whether disposition decisions improve when patients referred through second-stage (CDU) GK

Coefficient of interest: δ

Benefits of two-stage gatekeeping system

	No patient routed through CDU	All patients routed through CDU
Avoidable hospitalizations	5.23%	1.76%
Wrongful discharges	1.16%	0.67%

Avg. treatment effect (ATE)

- Avoidable hosp. = -3.5%
- Wrongful discharges = -0.49%

→ When patients routed through the two-stage system *both* errors go down

Benefits of two-stage gatekeeping system

	No patient routed through CDU	All patients routed through CDU
Avoidable hospitalizations	5.23%	1.76%
Wrongful discharges	1.16%	0.67%

Avg. treatment effect (**ATE**)

- Avoidable hosp. = -3.5%
- Wrongful discharges = -0.49%

→ When patients routed through the two-stage system *both* errors go down

Avg. treatment effect on the treated (**ATT**)

- Avoidable hosp. = -9.3%
- Wrongful discharges = -1.2%

→ ED physicians especially good at identifying patients who would benefit most from CDU second opinion

More ED or CDU capacity?

- **Counterfactual:** redeploy CDU capacity to the ED to reduce crowding
 - 1.5m patient hours in ED
 - 326,000 patient hours in CDU (=20% of ED hours)
- If there is no CDU, then adding 20% more capacity to ED reduces ED busyness by $\sim 0.6\sigma$:
 - leads to 0.14% reduction in avoidable hospitalizations (and slight increase in wrongful discharges)
- Keeping ED at its observed capacity and retain CDU:
 - leads to 1% reduction of avoidable hospitalizations
- **Summary:** the net effect of the CDU in the study hospital, after accounting for the opportunity cost of its resources, is a relative reduction of the avoidable hospitalization rate by 16.5%

The background is a green-tinted collage of various business school scenes. It includes a large crowd of students in a lecture hall, a modern glass-fronted building with 'INSEAD' signage, a group of students in a meeting room, and silhouettes of students in a gymnasium.

INSEAD

The Business School
for the World®

Europe

| Asia

| Middle East