Introduction:
We are working on an empirical study of how emergency department (ED) physicians balance the needs of the patients currently in treatment with the needs of those waiting to be treated. The fundamental trade-off that the physician faces is that devoting more time and effort to the patient at hand leads to higher quality care in expectation for that patient but delays care for all those waiting to be served. What makes this trade-off particularly difficult is that there is uncertainty on both sides of the decision. The physician does not know ex ante the value of ordering one more test for a current patient or how much a waiting patient is going to deteriorate while waiting. The average emergency department (ED) patient likely assumes that he or she will receive the same best clinical care regardless of external factors such as the busyness of the ED. However, we show that physicians adjust the amount of care given to patients as a function of crowding. We also show that reducing the quality of care for current patients can be locally harmful but socially optimal.

We conduct a detailed econometric analysis of the medical treatment delivered during more than 200,000 emergency department visits at a major U.S. teaching hospital. We observe the number of pathology lab tests and radiology imaging scans ordered for each patient and relate these counts to a set of covariates including chief complaint, triage level, and the number of patients in the waiting room. This research design allows us to make the following three contributions:

1. We show that doctors adjust work content per patient as a function of system load. For example, for abdominal pain patients, we find that the expected number of lab tests and scans drop by 15% and 13% respectively as the waiting room census moves from the 10th percentile to the 90th percentile.
2. We show that this reduction in work content comes from a reduction in the probability of receiving any orders, not from a simple leftward shift of the distribution. In other words, doctors are not dropping the marginal test for each patient. Rather, they are dropping the marginal patient from all testing! For example, the probability of an abdominal pain patient receiving no labs triples from 7% to 21% as the waiting room census moves from the 10th percentile to the 90th percentile. Conditional on being tested, the number of tests does not change with load.

3. We find no evidence that doctors are able to systematically select less sick patients to receive less diagnostic testing as doctors respond to load. This is seen in the fact that the probability of a lab test having an abnormal result does not change with load. This is counter to what one would see if doctors were able to systematically select the less sick patients to drop from testing, in which case the probability of an abnormal return would increase. Furthermore, we find that for some patients, the reduction in testing is associated with lower clinical quality. For example, for chest pain patients, the reduced testing leads to an increased chance of an ED revisit within 24 hours.

This study raises interesting managerial questions for emergency medicine physicians and hospital administrators. Primary among them is the question, “is this an acceptable way to deal with crowding in the ED?” While theory shows that trading off quality for speed can be rational and socially optimal, it is not necessarily so if the costs are incorrectly valued or if there are other managerial options, such as increasing staffing or testing capacity. Furthermore, doctors currently eliminate all testing for some patients, which provides a large improvement in speed but comes at a significant quality cost for those who are eliminated from testing. Perhaps reducing testing intensity would provide a more favorable speed/quality tradeoff.

We found the emergency department to be a useful setting in which to study the impact of crowding on work content and the two ways that work content can be endogenously controlled; through intensity truncation and patient truncation. Future research should explore if and how these behaviors manifest in other settings. One potential area is security screening at airports and borders. Here, the security inspector faces a similar dilemma of deciding how much time and effort to spend on the current customer while making other customers wait.