**Behavior-Aware Queueing: The Finite-Buffer Setting with Many Strategic Servers**

**ABSTRACT** - Service system design is often informed by queueing theory. Traditional queueing theory assumes that servers work at constant speeds. That is reasonable in computer science and manufacturing contexts. However, servers in service systems are people, and, in contrast to machines, systemic incentives created by design decisions influence their work speeds. We study how server work speed is affected by managerial decisions concerning (i) how many servers to staff and (ii) whether and when to turn away customers, in the context of many-server queues with finite or infinite buffers (M/M/N/k with \( k \in \mathbb{Z}^+ \cup \{\infty\} \)) in which the work speeds emerge as the solution to a noncooperative game.

We show that a symmetric equilibrium always exists in a loss system (N=k) and provide conditions for equilibrium existence in a single-server system (N=1). For the general M/M/N/k system, we provide a sufficient condition for the existence of a solution to the first-order condition and bounds on such a solution; however, showing that it is an equilibrium is challenging due to the existence of multiple local maxima in the utility function. Nevertheless, in an asymptotic regime in which demand becomes large, the utility function becomes concave, allowing us to characterize underloaded, critically loaded, and overloaded equilibria.

**SPEAKER BIO** – Amy Ward’s research focuses on the approximation and control of stochastic systems, with applications to the service industry. Much of her past work has focused on the impact of customer impatience and abandonments on performance. Her recent work investigates the interactions between behavioral incentives and operational efficiency in service systems. Please view her personal web site for a list of her published papers. Please view here for a recent Chicago Booth Review Q&A.

Ward is the Editor-in-Chief of the journal Operations Research Letters and serves as the Stochastic Models Area Editor for the journal Operations Research. In the past, she has been chair of the Applied Probability Society (term 11/2016-11/2018).

Prior to joining Booth, Ward was Professor of Data Sciences and Operations at the University of Southern California Marshall School of Business. She has also been a Visiting Associate Professor in the Computing and Mathematical Sciences Department at Cal Tech, and an Assistant Professor in Industrial and Systems Engineering at the Georgia Institute of Technology. Outside of academia, during her doctoral studies, she spent several summers at AT&T Laboratories.

Ward earned both a PhD and MA from Stanford University, and she holds a BA from Claremont McKenna College.