Research Statement

Chamsi Hssaine
https://chamsihssaine.github.io • ch822@cornell.edu

My research centers around **algorithm and incentive design for smart societal systems**. Combining rigorous theory, data-driven approaches, and large-scale optimization, my work develops tools for the *people-centric* design and operations of these systems. In particular, a unique and unifying thread across all my research is a focus on incorporating more realistic models of behavior under incentives, and better understanding the effects of policy decisions on all stakeholders. Much of my current research has tackled these questions within the context of *smart transit*. More broadly, I am interested in using a large methodological toolbox (leveraging techniques from game theory, applied probability, statistics, and optimization) to study how people collaborate, compete, learn from, and affect one another in a wide variety of socioeconomic systems.

1 Current research themes

Platforms like Lyft and Uber have forever changed urban mobility, and have led to a growing research community devoted to the design and operations of these platforms. In recent years, however, many have started to ask more critical questions regarding the broader societal context in which these platforms operate, and these questions have directly and indirectly influenced my research. My algorithmic work on smart transit has focused on *how ride-hailing services can be leveraged as a part of a larger transit ecosystem.* Exposure to issues in ride-hailing platforms and public transit has also led me to think more broadly about how more nuanced and realistic models of agent behavior affect platform operations and market outcomes. This is reflected in my other research thrusts on inter-firm competition, cooperation in multi-agent environments, and the impact of uncertainty on human behavior.

Smart transit operations. In recent years, large American cities such as Austin and Portland have announced plans to create *integrated mobility marketplaces* to revitalize ailing public transportation systems [9]. One feature of this new ecosystem is that mass transit schedules can be more dynamic, adapting to changing demand and exploiting ride-hailing for first/last-mile connections. In **Real-Time Approximate Routing for Smart Transit Systems** [4] (joint with Sid Banerjee, Noémie Périvier, and Samitha Samaranayake, **selected as a finalist for the 2021 INFORMS Minority Issues Forum Paper Competition**), we formulate the problem of demand-reactive transit network design at scale. We develop fast algorithms with constant-factor approximation guarantees under practical assumptions (pre-specified set of feasible bus routes, limited transfers), and show that without these assumptions, no constant-factor approximation is possible. We moreover demonstrate the *practical* efficacy of our approach on real-world datasets, vastly outperforming state-of-the-art ILP solvers for large problem instances.

Another central question in this setting is how to *price hybrid transit modes* to maximize welfare. This involves a nested optimization problem, with a transit agency choosing schedules/prices at the upper level, and commuters choosing utility-maximizing transit modes at the lower level. In **"Plan Your System and Price for Free: Fast Algorithms for Multimodal Transit Operations"** [3] (with Sid Banerjee, Qi Luo, and Samitha Samaranayake), we leverage linear programming duality to convert this nested problem into a single optimization problem. Our approach is much faster than existing approaches to transit network design

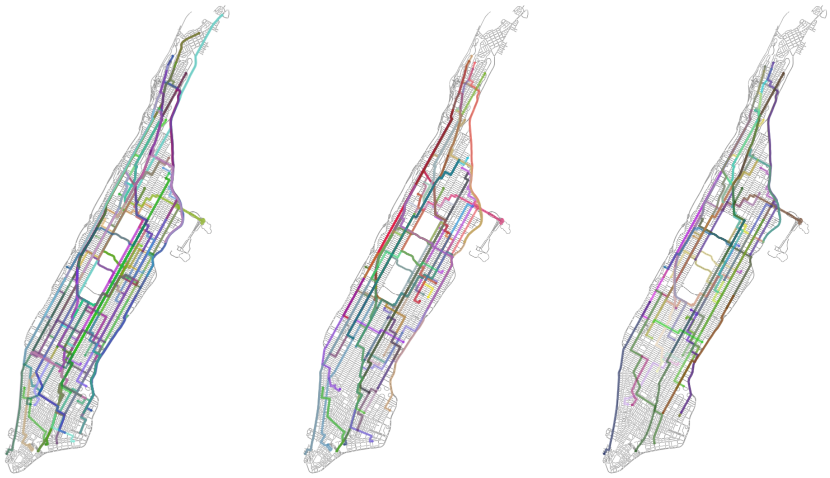

Figure 1: **Case study on the Manhattan road network: optimal mass transit routes for various commuter preferences** [3]

based on bi-level decomposition techniques. We demonstrate this by deploying it on real-world datasets to study a wide variety of questions, such as the welfare impacts of integrated mobility marketplaces and the effect of parameters such as ride-hailing costs and transfer frictions. Code and datasets for both projects are available at `https://github.com/SmartTransit-Cornell`.

Competition and cooperation. A hallmark of smart transit systems is the interplay of competition and cooperation between various stakeholders. This has led me to study the theoretical foundations of how data-driven decision-making impacts these two key facets of agent interactions.

A critical assumption in the extant literature on dynamic pricing and demand learning is that the firm is a *monopolist*. The problem of characterizing market outcomes when *competing* firms deploy popular online learning algorithms, however, has been explored to a far lesser extent, leaving open the possibility of unintended consequences when firms deploy these algorithms in real-world settings and treat their market environment as a black-box. In **"Pseudo-Competitive Games and Algorithmic Pricing"** [2] (with Sid Banerjee and Vijay Kamble), we consider games of price competition possessing a simple property: that *the revenue of a firm at a given price is independent of the prices of all firms setting strictly lower prices*. We demonstrate that this property is satisfied by many well-validated behavioral models involving the adaptive formation of consideration sets. Our work demonstrates that, under this property, price trajectories based on natural learning dynamics may converge to outcomes in which firms can experience unbounded losses in revenue compared to the best price equilibrium. To address this concern, we propose a novel learning algorithm which not only provably avoids convergence to such bad outcomes, but also successfully converges to this best equilibrium in numerical experiments.

In **"Information Signal Design for Incentivizing Team Formation"** [1] (with Sid Banerjee), we consider the opposite end of the spectrum of agent interactions. Motivated by a key challenge faced by massive open online courses, we study the use of Bayesian persuasion (also known as signaling) for organic, socially optimal team formation when agents have strong preferences over each other based on perceived aptitudes. Our contribution in this work is two-fold: (*i*) we develop tractable, asymptotically optimal policies for endogenous team formation, and (*ii*) we uncover the importance of agents' knowledge of their own abilities in determining the success of signaling mechanisms.

Impacts of algorithmic (un)certainty. My work on optimal signaling for team formation adds to a growing line of work on how a platform can use uncertainty to its advantage in the design of signaling mechanisms. In reality, however, there is a limit to how much uncertainty individuals are willing to tolerate in their experiences on a platform (be it as students, workers, or shoppers). My final thread of research explores the power and limitations of using uncertainty as an algorithmic tool in people-centric systems.

Critical to the success of a smart transit system that leverages ride-hailing is understanding how ride-hailing platforms compensate drivers. In **"Earning sans Learning: Noisy Decision-Making and Labor Supply on Gig Economy Platforms"** [5] (with Daniel Freund), we focus on two key features of platform compensation schemes: (*i*) that drivers have little insight into the algorithms that determine their earnings, and (*ii*) that driver earnings can be highly variable. An important corollary of this latter fact is that similar drivers may earn vastly different amounts for a similar amount of work, which calls into question the *fairness* of dynamic compensation schemes. Motivated by these observations, as well as our own empirical exploration of a publicly available ride-hailing dataset for the city of Austin, we study the problem of designing fair profit-maximizing compensation schemes for gig economy workers. We uncover a novel supply inefficiency that arises when earnings are volatile and hard to predict, and demonstrate how seemingly fair dynamic compensation policies can insidiously discriminate based on earnings. With an eye towards avoiding these undesirable outcomes, we develop a static compensation policy that is asymptotically optimal among all fair policies, and leverage this policy to derive analytical and numerical insights into the role of variability in gig worker compensation.

Another way in which uncertainty can be leveraged by online platforms is a practice known as *opaque selling* in which a customer allows a firm to choose the version of a product they receive. Recent work on the inventory benefits of this practice fail to account for an important fact: that uncertainty comes at a price. In particular, companies must offer these opaque products at a discount in order to incentivize customers to accept a possibly less-preferred item. In **"Timing Opaque Promotions,"** [6] (ongoing with Daniel Freund and Jiayu Zhao), we demonstrate that a little bit of uncertainty goes a long way in this respect. In particular, we show that exercising the opaque option infrequently, *and only very close to the end of the inventory cycle*, retains the inventory cost savings of exercising it without reserve, all the while minimizing the costs incurred from disrupting customers' purchase experiences.

2 Longer-term research agenda

In addition to further developing the above themes, my longer-term research agenda extends to other broad questions in the space of people-centric operations.

Repeated interactions and long-term impacts. A common assumption in the operations literature is that agents only interact with a firm once, or infrequently enough that previous interactions do not impact current decisions. It is well-known, however, that this assumption fails to hold in many real-world settings. Indeed, in both my industry internships at RAND and at Amazon, my projects involved building data-driven models to investigate the validity of this assumption in the public health and advertising sectors, respectively. At RAND, leveraging years' worth of longitudinal survey data, I explored the long-term impacts of individuals' and their social networks' prior vaccination experiences on current vaccination decisions. Similarly, as an applied science intern on Amazon's Personalization and Long-Term Optimization team, I investigated the effects of various A/B tests to understand the drivers of customers' long-term engagement with ads, producing actionable insights that the Sponsored Products team has incorporated into its ad placement toolbox.

In the future, I plan to leverage these experiences and explore this space of problems. One open direction in this regard is related to my work in [2] on unintended (and potentially disastrous) consequences of algorithmic pricing in smart societal systems. Specifically, though adequate exploration is crucial to guaranteeing strong performance of learning algorithms, if exploration results in customers experiencing highly variable prices, or, within the transportation context, being matched to trips with long detours, this will likely affect their trust in the marketplace and lead them to explore outside options instead. In another vein, in line with my work in [1] that considered a one-shot model of cooperation, I am interested in the theoretical underpinnings of incentivizing cooperation in repeated settings with little distributional information. How does learning affect mechanisms for cooperation?

Finally, my work in [5] is directly in line with this research theme. In particular, one key assumption is that gig economy workers are myopic: their decision to stay on the platform only depends on their most recent earnings. A more realistic model of behavior, however, would assume that workers use simple rules to *learn* the true distribution of rewards given their earnings history. Developing tractable models of such strategic behavior to derive insights into optimal compensation schemes is an open question that I am excited to study.

Intelligent infrastructure and smart societal systems. In the future, I will continue to pursue projects that work toward answering an overarching question that guides a part of my research: *How can we design a fair, efficient, and sustainable integrated transit ecosystem?*

From the ride-hailing platform's perspective, this involves being forward-looking in terms of regulations that local governments may impose in the future. For example, a 2016 investigation of a month's worth of Uber data in Washington D.C. found that neighborhoods with more people of color experienced longer wait times than whiter (more central) neighborhoods [8]. Such second-order effects may well become central concerns as, in 2015, the U.S. Supreme Court ruled within the context of the housing market that entities can be held accountable for this sort of unintentional discrimination – also known as "disparate impact" [7]. In the future, I plan to investigate how fairness constraints on metrics such as waiting times and passenger fares impact these services' matching and pricing operations.

From a public policy perspective, key to the vision of the smart city of the future – and to the success of an integrated mobility ecosystem in particular – is aligning the incentives of a number of stakeholders: the individual commuter, for whom a central concern is *coverage* (fairness); public transit providers, whose success metric is *ridership* (efficiency); and profit-maximizing ride-hailing firms. Specifically, one critical consideration is that of *contracting* with these platforms: *What is the appropriate timescale of contracting? What sort of service-level guarantees should form the basis of these contracts? Can these be implemented in an incentive-compatible way?* After my doctoral studies, I plan to investigate the fundamental limits and possibilities of mechanisms in this space. Importantly, I will continue to ground my research from a data-driven perspective, primarily relying on openly available public transit datasets. Moreover, through the National Science Foundation grant through which I am funded, I have built an interdisciplinary network of experts in online decision-making and control, transportation planning, and urban sociology. I will leverage this network to work toward my overarching goal of producing impactful research for large-scale urban systems.

References

[1] Siddhartha Banerjee and Chamsi Hssaine. Information signal design for incentivizing team formation. *Extended abstract in 14th Conference on Web and Internet Economics (WINE 2018)*, 2018.

[2] Siddhartha Banerjee, Chamsi Hssaine, and Vijay Kamble. Pseudo-competitive games and algorithmic pricing. Under review at *Management Science*.

[3] Siddhartha Banerjee, Chamsi Hssaine, Qi Luo, and Samitha Samaranayake. Plan your system and price for free: Fast algorithms for multimodal transit operations. Under review at *Operations Research*.

[4] Siddhartha Banerjee, Chamsi Hssaine, Noémie Périvier, and Samitha Samaranayake. Real-time approximate routing for smart transit systems. *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, 5(2):1–30, 2021.

[5] Daniel Freund and Chamsi Hssaine. Earning sans learning: Noisy decision-making and labor supply on gig economy platforms. Under review at *Management Science*.

[6] Daniel Freund, Chamsi Hssaine, and Jiayu Zhao. Timing opaque promotions. Working paper.

[7] NPR. In fair housing act case, supreme court backs 'disparate impact' claims, 2015.

[8] The Washington Post. Uber seems to offer better service in areas with more white people. That raises some tough questions., 2016.

[9] U.S. Department of Transportation. Smart City Challenge, `https://www.transportation.gov/smartcity`.