Research Statement

Changhwa Lee

October 18, 2021

I study various topics in microeconomic theory with a particular interest in mechanism design, information design, and rational inattention. My works are motivated by, but not restricted to, understanding how data is generated, used and monetized, and designing markets for and regulations over data. Inspired by this, I am currently working on two broad categories: markets for and regulation over data and discrimination.

I. Market for and Regulation over Data

My job market paper, "**Optimal Recommender System Design**," is motivated from an observation that consumers often rely on intermediaries' recommendations in deciding which product to purchase; but at the same time, sellers of products pay to the intermediaries in exchange for recommending their products. Intermediaries such as Amazon and Google have an immeasurable amount of consumer data from which they can infer which products are a good match for consumers. They use this information to give consumers recommendations, but at the same time, to collect payments from sellers. This is a major, but understudied, business model of monetizing data. I consider a problem of a monopolistic intermediary that designs a pair of a recommendations rule and a transfer rule to maximize the revenue it collects from sellers. An auction with bonuses and penalties implements the optimal recommender system: if products are good for the consumer, then sellers of the products get bonuses in that for them to win the auction to be recommended to the consumer, they can still win the auction even if their bids are somewhat lower than what their competitors bid. This is the key principle behind auction systems of real-world intermediaries such as Amazon and Google.

Using this characterization, I explore regulatory issues on these intermediaries. The main issue I explore is whether the intermediaries should be allowed to use data about sellers. When selling through intermediaries, sellers leave immeasurable data about their products, target consumers, sales, and margins. Intermediaries use sellers' data to raise more revenue is no secret that regulatory bodies suspect as anti-competitive. The key question to this concern is whether the data necessarily harms consumers and sellers. The answer I provide is that additional information does not necessarily harms consumers and sellers, and therefore, these regulations need to carefully think about benefits and harms of additional data before executing them. I provide sufficient conditions under which additional information is beneficial or harmful, and discuss what kind of additional information are beneficial or harmful. Beside the additional information, I also explore welfare-maximizing recommendations rule and find that, relative to the revenue-maximizing, it decreases the intermediary's revenue, but increases consumer surplus and sellers' profits. I hope to extend this framework to study further regulatory issues, including private-label products and competition in search engine services.

II. Discrimination

One of the main concerns in economics is to understand how discrimination arises and can be reduced. With the increasing availability of data, algorithms and predictive tools are proposed to reduce group disparities. For example, predictive policing algorithms are proposed to inform police where to monitor independent of their prior bias; hiring algorithms are proposed to decide whom to hire independently of genders. These algorithms, however, often resulted in widening the disparities: police monitoring neighborhoods of darker skins more often; hiring algorithms disproportionately favoring men over women. Central to this issue is how algorithms interact with human decision-makers. I use economic theory to understand the interactions and their implications on designing fair algorithms to reduce group disparities.

The marginal outcomes test (Becker (2010)) has become a 'go-to test' of taste-based discrimination in classification or allocation settings. In my paper, "**Outcome Test for Policies**" (with Mallesh Pai and Rakesh Vohra) argues that the test cannot be used in case of a policy, or an algorithm, and proposes an alternative statistical test to identify whether a principal designs the policy with discriminatory taste. We consider a model where a principal designs a policy, a mapping from signal (data) to decisions. The principal commits to the policy, which in turn affects on agents' incentive to take actions. In such environment, the marginal outcome test by principal has discriminatory taste. This is because the principal with commitment takes policies change agents' incentive to take actions into account when designing policy, the part that is not captured in the marginal outcome test. We propose a new statistical test which deviates away from the marginal outcome test to incorporate the incentive effects.

When designing a fair algorithm, the first and most important job is defining what a fair algorithm means. One dominant approach in computer science has been to require algorithms to have the same false positive and negative rates across groups. Another paper of mine "**Fair Prediction with Endogenous Behavior**" (with Christopher Jung, Sampath Kannan, Mallesh Pai, Aaron Roth and Rakesh Vohra) explores this question in a canonical setup where a principal designs a policy (algorithm) to minimize aggregate crime rates, where a policy is a mapping from signal (data) to decisions. When criminals' decisions to whether to commit a crime are endogenously determined in response to the algorithm, the principal has to think not only about whether a given individual is a criminal, but also about how his decision is going to impact on others' incentive to commit a crime. For a large class of signal distributions, the principal in fact only cares about minimizing incentive for individuals to commit a crime, which equalizes the size of incentives across groups. This result shows the limit of often-used criteria of equalizing false positive and negative rates and a new principle for designing fair algorithms using tools from economics.