Research Statement

Changhwa Lee

Department of Economics, University of Pennsylvania

I am interested in the economics of data, specifically how data is generated, used and monetized. I use the tools of mechanism and information design to understand how new data sources and their regulation affect incentives and discrimination. My current research belongs to two broad categories: market and regulation for data, and discrimination.

I. Market and Regulation for Data

My job market paper, "**Optimal Recommender System Design**," studies intermediaries like Amazon and Google that recommend products and services to consumers for which the intermediaries receive compensation from the recommended sellers. Nevertheless, consumers will find these recommendations useful if they are informative about the quality of the match between the sellers' offerings and consumers' needs. The intermediary would like the consumer to purchase the product from the seller who pays the most for a recommendation but is constrained because consumers will not follow the recommendation unless it is in their interest to do so. I frame the intermediary's problem as a mechanism design problem in which the mechanism designer cannot directly choose the outcome, but must encourage the consumer to choose the desired outcome. I show that in the optimal mechanism, the recommended seller has the largest non-negative virtual willingness to pay adjusted for the cost of persuasion. The optimal mechanism can be implemented via a handicap auction.

I use this model to examine the regulatory question of whether platforms should be allowed to use data reflecting sellers' private information, such as margins and bidding history. The use of data always benefits the intermediary, but can either benefit or harm the consumers and sellers. A special class of data is interpreted as the intermediary monopolizing a product market with private label products, and this is shown to benefit the consumer. I also examine a welfare-maximizing mechanism: relative to the revenue-maximizing mechanism, it reduces the intermediary's revenue but increases the consumer surplus and sellers' profits. An alternative interpretation of the model as a search engine is discussed. I intend to extend this framework to study additional regulatory issues, including consumer data privacy laws and regulation of targeted political advertising.

II. Discrimination

The ubiquity of data as well as the algorithms to analyze them, in principle, will allow organizations to make decisions or provide services in a race- or gender-neutral manner. Yet, there are a number of well-documented instances in which the deployment of such algorithmic tools have widened disparaties rather than narrowed them. Computer scientists have focused on algorithmic fixes to this issue. This, however, ignores the fact that algorithms interact human decision-makers. I use economic theory to understand the interactions and their implications for the design of fair algorithms to reduce disparities.

My paper "**Outcome Test for Policies**" (with Mallesh Pai and Rakesh Vohra) proposes a statistical test for identifying whether a policy or an algorithm is designed by a principal with discriminatory tastes. The test can be used for identifying, for example, whether predictive policing algorithms are discriminatory against minority neighborhoods. We also argue that the marginal outcome test (Becker (1993)), the most popular test of taste-based discrimination, fails for policies. We consider a canonical setup where the principal designs a policy (algorithm) that maps signals (data) to decisions for each group, such as whether to patrol or not for each area. The principal commits to the policy, which in turn affects agents' incentives to take action, such as whether to commit a crime. In this environment, the marginal outcome test fails because the principal not only cares about the marginal benefit of catching a criminal, but how patrolling changes agents' incentive to commit a crime. We propose a new statistical test that deviates from the marginal outcome test precisely as much as the incentive effect.

"**Fair Prediction with Endogenous Behavior**" (with Christopher Jung, Sampath Kannan, Mallesh Pai, Aaron Roth and Rakesh Vohra) explores the notions of fairness that policy makers should require algorithms to satisfy when interacting with human decision-makers. Our paper is the first to incorporate endogenous human behavior in the design of fair algorithms. In a simple setup where a principal designs an algorithm to minimize aggregate crime rates, we find that the most effective algorithm equalizes false positive and negative rates across groups. This result is in sharp contrast to the previous literature that abstracts away from the human-algorithm interaction and advocates equalizing thresholds on well-calibrated risk scores as the only policy that is both fair and effective. Together, these results demonstrate that it is essential to consider endogenous behavior in the design of fair algorithms - in other words, fair algorithms designed without taking human-algorithm interactions into account can be systematically misleading. One of our ultimate goals is to develop economics-based principles for the design of fair algorithms.