camila farres

Research Statement

Division of the Humanities and Social Sciences, California Institute of Technology
cfarresr@caltech.edu • camilafarres.com • Most recent version

My goal is to design personalized policies that help people by understanding their specific wants and needs. Similar problems may have different causes, and I believe that understanding people's unique characteristics is necessary to help them effectively. I refer to interventions that rely on individual characteristics as *personalization*. All of my projects relate to the challenge of designing personalization based on unobservable characteristics like preferences, biases, and beliefs. This challenge has implications for policy design and welfare.

The personalization framework

When we want to personalize based on an unobservable characteristic, we face many challenges. I think of personalization as a four-step pipeline. First, we need to define the objective and why we want to use personalization. For example, if a financial advisor wants to help their clients make better investment decisions, they try to understand the clients' risk preferences. Second, we need to measure the unobservable characteristics, like the clients' risk preferences. Third, we use what we learn about someone to offer them a personalized opportunity, such as a policy, a recommendation, an alternative, or a price. Examples of this are price discrimination and algorithmic recommendations, but also wage setting and medical diagnosis. Fourth, we want to evaluate whether personalization is more effective than a one-size-fits-all approach. For example, do students learn better with personalized education? Each step of personalization has its own challenges, and my research relates to some of them: what happens during measurement when people know about the personalization, whether the characteristic we measure is valid for personalization, and how to evaluate a personalized policy.

Do people change their behavior when they anticipate personalization?

The second step, measurement, has many challenges, but one that has been understudied is how people respond to measurement when they know about future personalization. This is the question I study in my job market paper, *Anticipating Personalization*. For example, a retailer learns a lot about us from our online behavior. If we expect them to give us better recommendations, then we would want to share as much about ourselves as possible, but if we expect them to charge us higher prices, then we would prefer they don't know us so well. To study this, I use a lab experiment that allows me to do three things: induce consumers' information (which is typically unobservable in the field), disclose all consequences of sharing the information (and test consumers' understanding of it), and randomize whether personalization is incentive compatible (i.e., whether it is in the consumers' best interest to share information accurately). I find that personalization does change the way consumers behave, but not in the way theory predicts. In particular, many consumers share information

accurately when they should not. This has welfare implications and suggests that policies that require institutions to disclose how they use consumers' information are not enough to protect consumers.

Are we measuring the relevant characteristics for personalization?

When we measure a trait for personalization, we should first check that it is the relevant trait for that objective. According to psychiatrists, people with ADHD have difficulty initiating and sustaining effort over time due to self-regulation problems. To help patients better self-regulate, we want to understand the source of the problem: is it because their preferences change or because there is a bias? In *Time Preferences and ADHD*, with Kota Saito and Junya Fujino's psychiatry lab in Japan, we measure three properties of time preferences: stationarity, time invariance, and time consistency, using effort-based convex time budgets over four weeks, and ask whether measured time preferences differ between patients with ADHD and neurotypical controls. The protocol paper was submitted, and data collection has started; it will continue through March 2027. We plan to extend the study to other mental health conditions such as autism, obsessive-compulsive disorder, depression, bipolar disorder, schizophrenia, and gaming disorder. The goal is to inform diagnostic methods and treatments for patients who have self-regulation problems.

Even if we know the right trait to measure, when we design the measurement instruments, we must choose which questions to ask or which data points to use. The paper *What Do We Really Know about Risk Preferences for Binary Lotteries?*, with Ted O'Donoghue and Charles Sprenger, currently under review, asks how our knowledge of risk preferences depends on the parameters we use to study them. Choices over binary lotteries have been used to document preferences inconsistent with Expected Utility (EU) and to motivate alternative models. However, most of the experiments use very similar parameters, so we provide a comprehensive exploration of the parameter space. For the narrow regions that are well studied, we find the same patterns. However, we find new patterns in the broader space that are inconsistent with EU and leading non-EU models. This shows that conclusions drawn from stylized examples can misrepresent preferences more broadly, which means we should be careful to use parameters that are relevant to the personalized intervention we are designing.

Is personalized education more effective?

My work in the field relates to the last step in the pipeline, evaluation. The Mexican government agency CONAFE provides public education in marginalized areas of the country. CONAFE developed and implemented a system known as tutorías, which tailors teaching to the interests and context of the student: tutors engage in one-on-one dialogue with students, and students reinforce what they learn by presenting it to their peers and then tutoring others. The tutoring model has expanded to countries like Thailand, Argentina, and Chile. However, it has never been rigorously evaluated. In *The Effects of Personalized Education in Marginalized Areas*, with Ana Lleó-Bono, we are designing an RCT to compare the effects

of tutorías on learning with those of standard educational approaches. We are currently designing an app with CONAFE to measure learning outcomes over time. The RCT will randomly assign communities to either standard education or tutorías. The goal is to inform educational policy given the high costs of this personalized program.

Looking Forward

My plan is to keep exploring the challenges of personalization to design better policies. With the increasing collection of personal data, this is more feasible than ever. However, there is much we do not understand about how to personalize and about its consequences. I will continue to work on the topic of my job market paper but move away from induced values to see how these results interact with some elements that my experimental design abstracts from. In particular, once we use personal information and people anticipate personalization, do they report who they are, who they want to be, or who they want others to think they are? For example, when a teacher asks a parent about their child's development, does the parent answer truthfully, say what they wish were true, or say what they want the teacher to believe is true? On the consequences of personalization, I want to understand how it changes social interactions. With AI, it could be possible to create content specific to each person. At what granularity should we personalize: to the common differences, or to the things that make us unique? Imagine if we all got completely different things; this could reduce what brings us together. Finally, what are people's preferences over personalization? By definition, personalization means that people get different things, but this could be perceived as unfair and reduce welfare.