Measuring Latent Behavioral Dynamics via Causally Constrained Generative Models

Application for the Robert Summers Fellowship on Economic Measurement

Yichao Jin, Ph.D.

School of Economic, Political and Policy Sciences

The University of Texas at Dallas

Richardson, TX 75080

Email: Yichao.Jin@UTDallas.edu

Portfolio: yichao2022.github.io

Date: May 2, 2026

Research Summary: Innovations in Behavioral Measurement

My research agenda focuses on advancing the frontier of economic measurement by bridging structural micro-econometrics with Generative Artificial Intelligence. Specifically, I develop novel methodologies to measure and forecast latent behavioral preferences (e.g., institutional trust, time discounting, and risk aversion) in data-sparse environments. My primary contribution to economic measurement is the development of the **Behavioral Digital Twin (BDT)** framework—a computational instrument designed to overcome the static limitations of traditional survey data by generating high-fidelity, causally constrained synthetic populations for counterfactual policy evaluation.

1. THE MEASUREMENT CHALLENGE: CAUSAL HALLUCINATIONS IN LLMS

Traditional economic measurement relies heavily on observational data or Discrete Choice Experiments (DCEs). While these methods are rigorous, they struggle to dynamically extrapolate out-of-sample behaviors under unprecedented policy shocks. Recently, Large Language Models (LLMs) have been proposed as "synthetic respondents." However, my research identifies a critical measurement failure in unconstrained generative agents: *causal hallucinations*. When tasked with simulating economic decisions, base LLMs frequently violate fundamental economic axioms, such as monotonicity (e.g., simulating increased vaccine uptake in response to longer waiting times). To utilize generative models for valid economic measurement, their outputs must be mathematically tethered to empirical reality.

2. THE BDT FRAMEWORK: A NEW INSTRUMENT FOR ECONOMIC MEASUREMENT

To resolve this, I engineered a three-layer architecture that redefines how we measure and project population-level behavioral responses. This framework serves as a novel data product, transforming static DCE data into interactive, dynamic simulation engines:

- **Layer 1: Neural-Behavioral Ingestion.** This layer measures the "Behavioral DNA" of empirical subjects by extracting latent parameters from stated-preference surveys and DCE data.
- **Layer 2: Probabilistic Decision Computing.** I embed Structural Causal Models (SCMs) directly into the agent's generative pathway. Instead of relying on stochastic text generation, the agent's utility function is constrained by econometrically derived parameters.
- **Layer 3: Synthetic Simulation Sandbox.** The fully aligned agents are deployed in a localized, Monte Carlo simulation environment to conduct stress-tests of counterfactual policies.

3. EMPIRICAL VALIDATION: MEASURING TRUST AND TEMPORAL FRICTION

I validated the measurement capacity of the BDT framework using a high-resolution DCE dataset on vaccination timing (n=1,027) collected in Wuhan, China. The empirical analysis measured the joint

elasticity of waiting time and institutional trust, revealing non-linear decision distortions caused by trust deficits.

When this empirical baseline was used to constrain generative agents, the BDT framework achieved a **34% reduction in frontier-deviation Mean Squared Error (MSE)** compared to unconstrained LLM baselines. Furthermore, the causal hallucination rate dropped from 50.0% to 22.2% in synthetic validation. The remaining hallucinations concentrate in data-sparse tail scenarios, which can be systematically suppressed through active-learning DCE designs that iteratively probe these low-density regions. This demonstrates that BDTs can measure and replicate the micro-foundations of human choice with a degree of precision suitable for rigorous economic forecasting.

4. COMPUTATIONAL RIGOR AND DATA SOVEREIGNTY

A critical challenge in modern economic measurement involving health or financial data is privacy compliance. To address the constraints of HIPAA and GDPR, the BDT data product is engineered to operate entirely on workstation-class hardware (e.g., Apple Silicon M4 Max architecture). By running 72B parameter models in a fully offline, local sandbox, researchers can synthesize highly sensitive behavioral data without exposing empirical training sets to commercial cloud APIs.

5. ALIGNMENT WITH NBER SUMMERS FELLOWSHIP

Having developed this framework from initial econometric specification through to empirical validation and synthetic deployment, I intend to use the NBER Summer Institute as a platform to refine its broader applications. The Conference on Research in Income and Wealth (CRIW) places a premium on innovative data products. My framework offers a scalable, causally robust instrument for measuring behavioral heterogeneity that traditional administrative data often misses. Participation would provide an invaluable opportunity to explore how causally constrained synthetic data can be standardized for applications in labor, health, and public economics. My working papers and the technical architecture of the BDT framework are available on my portfolio (yichao2022.github.io).