Contact information
Name: Michael Caosun
Institution and Department: Information Technology, MIT
Mailing address: 100 Landsdowne Street, Apt 1310, Cambridge, MA 02139
Email: mcaosun@mit.edu

Graduate Program and Study

I am a Ph.D. candidate in Information Technology at MIT. I have completed more than two years of Ph.D. coursework, including advanced sequences in microeconomic theory, econometrics, and machine learning, and will have completed coursework in industrial organization.

My tentative dissertation title is: **Essays on the Practical Application of LLMs: Negotiation, Skill Dynamics, and Organizational Performance.**

My tentative dissertation committee is: **Sinan Aral** and **Dean Eckles**.

Research interests and current projects

My research can be divided into two streams of work on LLM-based negotiation, and AI-driven productivity. One examines the design of LLM agents and tools for users and platforms, especially in negotiation settings. For example, I study how negotiation theory can inform the design of LLM-based agents on platforms. The other studies how organizations should use LLM tools. For example, when there is a trade-off between short-run productivity and long-run human skill through a mechanism like skill atrophy, how might that affect the optimal usage of these tools.

These themes appear in my published paper on AI negotiations and in two ongoing projects that will form a significant part of my dissertation.

Manuscript: ***Advancing AI Negotiations: New Theory and Evidence from a Large-Scale Autonomous Negotiations Competition*** **(Revise & Resubmit at PNAS)**

My existing research studies LLM-based negotiation agents at scale. We organized an international AI negotiation competition in which participants designed and refined prompts and strategies for LLM-driven agents. Using these agents, we conducted over one hundred thousand autonomous negotiations across scenarios that varied in structure and objectives, generating a large corpus of bargaining transcripts and outcomes.

The analysis shows that core principles from human-human negotiation carry over to AI-AI settings: agents that exhibit "warmth" in their language (positivity, gratitude, question-asking) reach agreements more often and, in integrative settings, tend to create

and share more value. More "dominant" agents are better at claiming value when deals occur, but they generate more impasses. We also document AI-specific tactics such as chain-of-thought prompting, prompt injection, and strategic concealment, and show that the best-performing agents combine classic negotiation principles with these AI-specific techniques. The paper makes the case for a theory of "AI negotiations" that links traditional negotiation research to the technical affordances and failure modes of autonomous agents in digital markets.

Ongoing project 1: The AI Augmentation Trap and LLM-induced Skill Atrophy

My second major project is a theory-driven and empirically informed analysis of how LLM usage affects worker productivity and long-run skill. It consists of a main theory paper and a shorter empirical validation exercise.

In *The AI Augmentation Trap: Over-reliance on AI Tools Under a Productivity–Skill Trade-off*, I develop a dynamic model in which organizations choose how intensively workers use LLM tools. LLM usage raises output through two channels: a skill-neutral gain (direct automation of tasks) and a skill-complementary gain that scales with worker capability (amplifying judgment-intensive work). At the same time, usage displaces practice and accelerates skill atrophy, modeled via a learning–forgetting process.

The model maps out a five-region policy space, depending on LLM and industry characteristics, that describes when LLM adoption leaves organizations better or worse off in the long run. In one region, firms optimally choose interior augmentation that raises short-run productivity but leaves them worse off in steady state than if they had never adopted LLMs at all—the "augmentation trap." This framework helps interpret heterogeneous treatment effects from recent LLM field experiments and speaks to how firms should design AI usage policies when human capital dynamics matter.

As a first empirical step, I have begun a simple external validation exercise that relates occupation-level AI exposure indices to employment growth, as a coarse check on the model's automation channel. Going forward, I hope to deepen this empirical component using richer microdata on LLM adoption and worker performance. One reason I would like to attend the tutorial is to see how other researchers are validating related models and to learn about promising data sets and designs.

Ongoing project 2: Designing LLM Negotiators based on Interaction-Aware Adaptation

My first ongoing project builds directly on LLM-based negotiations. The goal is to understand how negotiation tactics interact in AI-mediated bargaining and to design an adaptive negotiating agent that can systematically exploit those interaction effects.

I construct controlled negotiation environments (e.g., multi-issue bargaining with deadlines and outside options) and run round-robin tournaments between many fixed-tactic agents. Because LLM agents are programmable and reproducible, this approach lets me estimate both the marginal effects of individual tactics and the interaction effects among them.

Using these estimated interaction patterns, I am developing an adaptive agent that learns which tactic sequences perform well against different types of counterparts. The agent observes negotiation messages, updates beliefs about the counterpart, and adjusts its tactics in real time. In preliminary experiments, this design both creates and claims more value than any single fixed-prompt agent.

The project is aimed at understanding how to design agents that create value rather than only claim it, which is especially relevant for platforms that earn revenue from the value created in transactions and need to balance the needs of multiple parties.

Fit with the Digital Economics and AI Tutorial

I think my present focus on applications of LLMs in negotiation and work make me a good fit for the tutorial. My past and current projects are closely aligned with the program's focus on digital markets and AI. The negotiation work contributes to understanding how LLMs can reshape bargaining and market design on platforms; the augmentation-trap work connects to debates on AI, productivity, and long-run human capital in digital work.

Methodologically, I combine formal modeling, LLM-based experiments and simulations, and empirical analysis using occupation-level data, and I am in the process of moving toward richer microdata. I would benefit from feedback on both identification and design choices from researchers working on related questions in digital economics and AI.

I am at a stage in my research career where meeting a broader group of scholars at the NBER Digital Economics and AI tutorial could have a large impact on my dissertation and future work. Talking to other scholars about how the space is evolving would help me sharpen my questions and better situate my work within the broader literature on digital platforms and AI.