Contact Information:
Mariya Pominova
Phone Number: (339)-223-1123
Email: mariya.pominova@duke.edu

Description of research interests and ongoing research:
I am an Economics Ph.D. student at Duke University and my main areas of interest are macroeconomics and labor economics. Specifically, I am interested in studying the determinants and macroeconomic implications of labor inequality. Artificial intelligence has permeated so many aspects of day-to-day in the past decade. Algorithms are being used every day to determine credit worthiness, market employment opportunities, products, and services, and assist with hiring decisions. I am also interested in studying how artificial intelligence in recruiting and hiring practices impacts labor inequality and what implications this has on existing policies that are intended to increase workplace diversity.

My interest in understanding implications of artificial intelligence on societal outcomes stems from a project I worked on as a research assistant at the Federal Reserve Board of Governors. For this project, we were interested in understanding how credit covenants in lending agreements can be utilized as an early indicator of changes in financial cycles. I developed a mechanism that could sift through millions of SEC filings to identify and filter out credit agreements from within the filings using a Random Forest machine learning algorithm. The task we needed the model to do was akin to finding a needle in a haystack. Credit agreements represented less than 1 percent of SEC filings and appeared frequently as attachments to filings that weren't titled descriptively. Thus, to create our final model that achieved over 95% accuracy, we needed a very large training set with extensive examples of various credit agreements. Through this experience, I began to think about the impact of training set composition on algorithm effectiveness and the impact this might have when machine learning is used for selection of job candidates in hiring models.

For my third-year paper, I am working on incorporating artificial intelligence resume filtering into the baseline search and matching model to understand how full adoption of this technology will impact labor inequality. Nearly every fortune 500 company uses artificial intelligence (AI) resume screening as a part of their hiring process and survey evidence has shown that adoption of AI screening tools will continue to rise. While the use of this technology is actively becoming more widespread, the economics literature on questions pertaining to understanding, measuring, and regulating bias in hiring algorithms is still in its nascent stage. My research introduces a structural methodology for modeling how underrepresentation of marginalized groups can result in an algorithmic bias against hiring said group. In my paper, I prove that underrepresentation of a group in the training set will result in under-selection of this group by the algorithm when the firm is risk averse. The risk aversion in my model comes from firms being heterogenous in production technology and having a threshold productivity requirement for their workers in order to produce. Next, I use the Diamond-Mortensen-Pissarides model to simulate outcomes for various levels of AI integration into the firm-worker matching process. I will compare group differences in lifetime earnings, steady state distribution of employed workers and other relevant metrics across models with fully-human, fully-AI, and combined hiring operations methods of firm-worker matching

Additionally, I am working on an extension to my paper titled, The Pitfalls of Using Location Quotients to Identify Clusters and Represent Industry Specialization in Small Regions", which was published in *Review of Regional Studies* in 2024. This is a coauthored project with my advisors, Andrew Crawley and Todd Gabe, which began as my master's thesis. In this paper, we proposed a method for evaluating the reliability of the location quotient – a prevalent measure of regional industrial specialization – at a given level of aggregation by adding a single hypothetical establishment to each industry in the region and assessing the squared difference between the original and augmented measures. The extension I am working on with Andrew Crawley and Christopher Shultz is an adjustment to the location quotient which would allow us to more accurately capture the specialization in very small rural areas. This extension is useful because in small rural areas, investment funding to support industries often is provided to the most specialized industries. However, one of the most common measures of specialization is often not useable in these areas.

I would like to attend the NBER AI Tutorial in order to meet and learn from other economists that are modeling implications of artificial intelligence on inequality. There is not an extensive literature on my research question, and learning from researchers studying artificial intelligence can help me further my research agenda. Additionally, I would like to meet other Ph.D. students interested in similar topics to build my network of collaborators for the future.

Description of previous graduate coursework:
I am currently a second-year student in the Duke Economics Ph.D. program. However, by the time this tutorial takes place I will be most of the way through my second year. I believe my research on my third-year paper is in a mature enough place for me to immensely benefit from attending this tutorial and the subsequent NBER Digital Economics and AI meetings.

By the time of this tutorial, I will have completed the following field courses:

- *Economics of Race and Racism:* This course covered structural modeling techniques for identifying and measuring discrimination in labor markets, housing markets, education, and criminal justice system.

- *Macro-Labor Theory and Application:* This course covered theory and extensions of the canonical search model and how these can be used to study labor market power, discrimination in the labor market, wage-price-spirals, and unemployment scarring.

- *Numerical Methods in Macro:* This course covered computational methods in macroeconomics. We discussed techniques for numerically solving models in discrete and continuous time, such as the endogenous grid method, and estimation methods using maximum likelihood and generalized methods of moments.

- *Empirical Models of Matching and Sorting:* This course provided an overview of modeling techniques for answering questions with two-sided matching and endogenous sorting. We discussed applications in residential segregation, urban agglomeration, school choice, and labor markets.

- *Expectations in Macroeconomics (Spring 2024):* This course covers the consequence of departures from full-information rational expectations for macroeconomic modeling. We will discuss topics including higher-order beliefs, signal extraction from endogenous variables, rational inattention, and imperfect memory.

- *Firm Heterogeneity and Macroeconomics (Spring 2024):* This course covers firm-level heterogeneity underpinning macro developments. We will cover topics on firm productivity, misallocation and aggregate growth, networks, and uncertainty.

Additionally, in my first year at Duke, I completed the following courses as part of the economics core sequence:

- *Macroeconomic Analysis I:* In the first half of this course, we covered the neoclassical growth model. In the later half, we covered Real Business Cycle models and asset pricing.

- *Macroeconomic Analysis II:* In the first half, we covered search and matching, heterogeneous agent models, and fiscal policy in dynamic general equilibrium models. In the second half of the course, we covered monetary policy theory.

- Microeconomics I: The first half of this semester covered decision theory concepts. The second half of the course we focused on monotone comparative statics.

- *Game Theory:* In this course, we covered the key game theoretical concepts such as dynamic games, repeated games, Bayesian updating, signaling models, etc.

- *Econometrics I:* This course covered core econometric methods such as regression analysis, causal inference techniques, instrumental variables, panel data, and approaches to handling missing and censored data.

- Econometrics II: In the first half of the course, we covered maximum likelihood, generalized method of moments, and panel data estimator theory. In the second half, we covered general time series econometrics theory.