MARTA BOCZOŃ

Research Statement

Nothing in life is to be feared, it is only to be understood. Now is the time to understand more, so that we may fear less.—Marie Skłodowska Curie

Everything starts with a question. Then, comes the execution. In my case, I was first exposed to ways of answering a research question before I learned how to formulate one. During my master's program in Statistics I learned about various statistical tools and estimation techniques used for cross-sectional, panel, and time series data analysis. Then, as a PhD student in Economics, I learned what I believe to be the essence of economic studies and namely, how to decide on a research question. For the past five years, I have been exposed to a variety of topics during seminars, brown bags, and job market talks, which I regularly attend, for a number of fields: Applied Microeconomics, Econometrics, Macroeconomics, and Experimental Economics. This is how I became a question-driven researcher passionate about each and every one of my projects. When deciding on a research question, I pursue topics that are of interest to me and discuss relevant current affairs. In particular, I never decide not to conduct an analysis because of initial unfamiliarity with either the topic, data or estimation. I believe that conducting research is a learning process that nobody should be afraid of; it is an exciting journey with us, researchers, in a driver's seat, deciding on whether or not to start the car and how far we want to go.

One of the reasons that I came to the University of Pittsburgh is that I was excited to work with my current supervisor, Professor Jean-François Richard, on one of his projects titled "Balanced growth approach to tracking recessions." After several months of working together, I was formally offered to be my supervisor's co-author—a result of offering worthwhile suggestions regarding methodology and proving great empirical skills. In this project (published in Econometrics) we propose a generic procedure to transform structural macroeconomic models into hybrid versions thereof in a way that preserves their policy relevance while significantly improving their recession tracking performance. As such, we illustrate that it is possible to achieve both empirical and theoretical coherence without the need to sacrifice one over the other. During the past three years I spent working on this project, I enjoyed the most frequent and stimulating discussions with my supervisor Richard, who I consider an extraordinary researcher. In addition to our joint publication, we are planning to continue our analysis by extending our sample period to the first quarter of 2021 to account for the ongoing pandemic triggered by COVID-19. Our goal is to find whether our proposed hybrid approach can track the ongoing recession to which there is no precedence. Moreover, since our model allows for policy interventions, we hope to be able to determine what policy, and on what timeline, would help to reduce the negative impact of the COVID-19 on the economy, and specifically, rising unemployment rates.

My interest in biology and medicine led me to work on a research project titled "Semiparameteric estimation of the relationship between recessions and health" (currently submitted for publication). After reading several papers on this topic, including a seminal work by Christopher Ruhm "Are recessions good for your health?", I noticed that the existing literature usually relies on a single proxy for economic conditions (mean unemployment rate) and assumes linearity of the underlying relationship. Therefore, after collecting qualitative evidence suggesting that mean unemployment rate fails to comprehensively characterize the current state of the economy, and that the relationship in question may be nonlinear, I addressed these two potential shortcomings. First, I introduced an additional proxy for economic conditions (volatility of unemployment rate) and second, I estimated

the relationship in question semi-parametrically in order to allow for any form of nonlinearity. In this project, I found that the relationship between economic conditions and health is not only non-linear but also statistically significant. I believe that this finding can potentially give rise to a new area of research that will lead to a better understanding of the mechanisms necessary for this relationship to exist in the first place.

On the recommendation of my supervisor, I was offered a co-authorship by Professor Alistair J. Wilson on a project titled "Goals, constraints, and transparent assignment: A field study of the UEFA Champions League." In this paper, we assess properties of an existing matching mechanism, which is a public drawing of football-team pairs in the Union of European Football Association's (UEFA) Champions League (UCL). Since the UCL is one of the most successful pan-European ventures, and certainly the one that receives the most enthusiasm from the general public, our study focuses on a widely-known mechanism, where participants' expected outcomes can shift by millions of euro depending on a particular realization of the draw. As a fan of European football myself, I became immediately interested in the topic. Parts of the project that I have enjoyed working on the most were data collection where I constructed a data set of scorlines from all UCL games in the last twenty years, learning about and formalizing the current UEFA draw mechanism, and estimating a structural model to assess each team's attacking and defending performance. While starting off the project, my co-author and I believed that the UEFA mechanism is non-optimal in terms of fairness, which we defined in terms of equal matching probabilities among unconstrained teams. However, a few weeks into the project, we found that the mechanism in question is virtually optimal and as such, cannot be improved upon without significantly complicating the current draw procedure, whose advantages are its transparency and simplicity. At this point in time, my professor and I took the nearest exit and since then, are continuing our journey heading in a slightly different direction. Specifically, in the process of revising the paper for a re-submission (R&R) at Management Science, we aim to demonstrate how the UEFA procedure can be useful in situations like assigning faculty members to offices or shifts to workers, assuming zero or one preferences over offices/shifts.

While working with the Survey of Consumer Finances (SCF) for one of my class projects, I stumbled upon the topic of my job market paper , which studies uncertainties in estimates of top income and wealth inequality in the United States. Specifically, I observed that the standard errors on the SCF estimates of households' average asset holdings are by no means marginal. This finding led me to review the related literature, which eventually helped me realize that the issue of data-driven errors (such as the SCF wide confidence intervals) is often not addressed in macroeconomic studies. Therefore, I decided to pursue this research topic and investigate to what extent empirical and structural macroeconomic analysis are vulnerable to data-driven errors. In order to obtain firsthand information about the SCF and discuss the project with those who run the survey, I applied for the Summer Dissertation Fellowship at the Board of Governors of the Federal Reserve System. I consider the three-and-a-half months spent at the Board as one of the most important milestones of my research career. I was given excellent mentoring, IT support, and access to the Federal Reserve Board Research Library. Thanks to shared research interests, complementarities, and discussions about possible co-authorship following my graduation, my work relationships with a number of Board members continue until this day. During my fellowship, I formulated my job market paper research question, and decided to conduct my analysis in the context of income and wealth inequality. I find that when analyzing income shares of the top 0.1 and 0.01 percent, the SCF becomes unreliable—a result of insufficient number of observations above the 99.9 and 99.99 income fractiles. However, when investigating less granular income shares of the top 10, 5, 1, and 0.5 percent, the SCF proves as reliable as administrative data in regard to estimating long-term trends in the observed income inequality and assessing their statistical significance. As of now I am investigating the impact of two sources of error: sampling and nonresponse. In my next project, I plan on estimating an additional source of error, measurement error. The reason for estimating this type of error stems from the qualitative evidence that measurement error may be substantial in self-reported financial surveys such as the SCF. Moreover, according to my research, measurement error is—as of now—the only unaccounted source of error that may bias the estimates and substantially widen confidence intervals.

Over the past five years I have transitioned from a student to a researcher, and developed a unique research style which helps me conduct my projects in the most efficient manner. Within a few weeks into a project, I start with a first draft. Having an early version of my paper not only allows me

to apply for conferences and get early feedback from my advisers, but also helps me to organize my thoughts and further structure the paper. Moreover, I always invest time at the beginning of a project to understand the data I am working with. In addition to tabulating and visualizing them, I learn about their deficiencies that may bias my outcomes and limit their external validity. Finally, I always use the software that is best-suited to answer a specific research question, regardless if I am familiar with it at first or not. During my PhD studies, I have worked with Fortran, Matlab, R, Stata, and Mathematica, each time choosing a software that best suits my research question and/or my coauthor's preferences. Specifically, I commonly use multiple software in one project, playing on each software strengths. For example, I often do initial data work in Stata, what allows me to efficiently analyze large data sets with hundreds of thousands of observations using build-in functions designed to tabulate, summarize, and visualize the data. Next, I conduct the estimation in Matlab, since it defines variables as multi-dimensional arrays—an ideal data structure for estimating econometric models. Finally, when conducting a Monte Carlo simulation exercise, I use Fortran with the objective of substantially reducing the computation time, often by a factor of ten or more. All in all, relying upon this different software allows me to have more control over the road I am on, and specifically, gives me the opportunity to choose highways over unpaved paths with potholes and strict speed limits.

To conclude, conducting research is a way to conquer the fear of unknown, broaden one's horizons, and challenge ourselves as well as the status quo. This is why I became a researcher, since "nothing in life is to be feared, it is only to be understood. Now is the time to understand more, so that we may fear less" (Marie Skłodowska Curie). Therefore, I strongly believe that I have much to offer to a research group like yours through interactions with colleagues, constructive participation to meetings and seminars across areas of research, a wide range of empirical skills, and a willingness to challenge conventional wisdom.