The feeling of the age

A quantitative analysis of the correlation between novelistic and economic sentiment

Research Proposal

Laura Gáti[*]

Macroeconomic research has in recent years awoken to the importance of expectations for macroeconomic fluctuations. The literature has focused on two main ways in which beliefs influence business cyles. The first questions the assumption of full information. Examples include Mankiw and Reis (2002), Sims (2003), Mackowiak and Wiederholt (2009) and Angeletos and Pavan (2009). The second, following Coibion and Gorodnichenko (2015), instead investigates alternatives to a rational expectation formation process.

An assumption implicit in the majority of this literature is that the interplay between beliefs and economic fluctuations takes place at business cycle frequencies. As such, the low frequency comovement properties of expectations and economic developments is less well understood. In this project, joint with literature scholar Daniella Gáti, we combine the insight of Alexopoulos (2011) that text is a useful source of information for long-run economic phenomena with Hoff and Stiglitz (2015)'s emphasis on the interplay between economics and culture. We ask: what predictive information do works of fiction contain for long-run economic fluctuations?

The project consists of a two-step procedure. First, we conduct a sentiment analysis of prize-winning novels of the Pulitzer, National Book Award, PEN Faulkner, PEN Hemingway and Book Critics Circle awards from the years 1948-2018. We assign the novels quantitative sentiment scores. Second, we investigate the correlation between the sentiment scores and macroeconomic aggregates such as GDP, total factor productivity, R&D expenditures, and so on. Once we have identified the frequencies where most of the comovement occurs via frequency domain methods, we explore the cointegration relationships and the effects of innovations to literary sentiment using a structural Vector

---
[*]Boston College, *Email*: gati@bc.edu, *Website*: https://sites.google.com/view/lauragati

Error Correction model (VECM). We hypothesize that while novels may not function as predictors of economic developments in the short run due to the time it takes to write and publish them, they do become valuable descriptors of the structural conditions of the economy at a particular time, as well as of medium-run optimism/pessimism regarding the economy.

The first semester of the Harvard Academy Scholars Program will be devoted to the compilation and cleaning of the dataset. Having received the Brandeis Provost Dissertation Fellowship, we have the funds to acquire e-book editions of all the award-winning novels from the above listed book awards for the time period 1948-2018. This vast dataset needs to be processed to be suitable for text analysis. The word embedding analysis will then be carried out in the second semester. This involves the development of algorithms that parse the textual data to extract sentiment scores, which in turn will form a time series of novelistic sentiment.

The third semester will then be spent with the economic part of the analysis. I plan to allocate two months to investigating the joint spectrogram of the novelistic sentiment series with the time series of macroeconomic aggregates. The remaining months of the semester are allocated to developing the suitable VECM framework and running initial reduced-form regressions. At this stage, I anticipate presenting the project at Harvard, and applying to future conferences. In the fourth semester, then, we will conduct the structural VECM analysis, exploring various identification strategies and robustness checks. This time will also be allotted to composing a first draft of the paper, presenting at various conferences and, if possible, presenting a second time at Harvard. Lastly, this period of time will be dedicated to writing the paper, and submitting it for publication. Revisions will also be carried out as necessary.

Throughout the time spent on working on this project, I also plan to develop my older projects. The first priority will be the publication of my job market paper. I anticipate that one third of my time in the Harvard Academy Scholars Program position will be spent on submitting and revising my job market paper. Conditional on time remaining, I intend to submit my second and third dissertation chapters for publication as well. From this perspective, I hope to present also these projects at Harvard.

For these reasons, I am applying to the Harvard Academy Scholars Program to support this important research on the interaction between macroeconomics and novelistic sentiment on the one hand, and my broad research on expectations in macroeconomics on the other. The exchange with the intellectual community at Harvard would further not only my proposed project, but through the development of my previous research, it would strengthen my entire research agenda.

References

Alexopoulos, M. (2011, June). Read all about it!! What happens following a technology shock? *American Economic Review 101*(4), 1144–79.

Angeletos, G.-M. and A. Pavan (2009). Policy with dispersed information. *Journal of the European Economic Association 7*(1), 11–60.

Coibion, O. and Y. Gorodnichenko (2015). Information rigidity and the expectations formation process: A simple framework and new facts. *American Economic Review 105*(8), 2644–78.

Hoff, K. and J. E. Stiglitz (2015, December). Striving for balance in economics: Towards a theory of the social determination of behavior. Working Paper 21823, National Bureau of Economic Research.

Mackowiak, B. and M. Wiederholt (2009). Optimal sticky prices under rational inattention. *American Economic Review 99*(3), 769–803.

Mankiw, N. G. and R. Reis (2002). Sticky information versus sticky prices: a proposal to replace the New Keynesian Phillips curve. *The Quarterly Journal of Economics 117*(4), 1295–1328.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics 50*(3), 665 – 690. Swiss National Bank/Study Center Gerzensee Conference on Monetary Policy under Incomplete Information.