Heterogeneity in Macroeconomics: Implications for Monetary Policy An Overview

Sofía Bauducco
Central Bank of Chile

Andrés Fernández
International Monetary Fund

Giovanni L. Violante
Princeton University
Centre for Economic Policy Research
Institute for Fiscal Studies
National Bureau of Economic Research

This volume collects some of the papers presented at the XXV Annual Conference of the Central Bank of Chile, which took place in November 2022 in Santiago, Chile.[1] The theme of the conference was **Heterogeneity in Macroeconomics: Implications for Monetary Policy**. The main objective of this conference was to invite some of the most prominent macroeconomists working on models with salient heterogeneity among households, firms, and banks to present their research and discuss how this class of models can inform the design of monetary policy.

The rapid surge in interest, among academic researchers and policymakers, in the nexus between heterogeneity and monetary policy is associated with the emergence of a new class of models referred to as HANK, an acronym for Hetergoeneous-Agent New Keynesian. HANK models combine two long-standing traditions of macroeconomic theory: (i) the new Keynesian approach to the study of business cycles and stabilization policies and (ii) the heterogeneous-agent incomplete-market approach to the study of the wealth distribution and of those

1. The full program is available on the Central Bank's website.

Heterogeneity in Macroeconomics: Implications for Monetary Policy, edited by Sofía Bauducco, Andrés Fernández, and Giovanni L. Violante, Santiago, Chile.

policies that offer social insurance, promote income mobility, and redistribute across households.[2]

The production and monetary policy blocks of this model are exactly the same as in the representative-agent new Keynesian (RANK) model and, as in that framework, they are summarized by three aggregate equations: (i) the Phillips curve, which specifies a relation between inflation and output dynamics; (ii) the Taylor rule, which summarizes how the monetary authority operates its main policy instrument, the nominal interest rate; (iii) and the Fisher equation, which links the real interest rate, the policy rate, and expected inflation. The crucial innovation lies in replacing the representative consumer, and hence the aggregate Euler equation (or the IS curve), with the modern theory of consumption and saving. The starting point of this theory is that households differ ex ante because of innate heterogeneity, and ex post because of idiosyncratic income shocks; then, due to financial market imperfections, these differences transmit to consumption, saving, and welfare. In equilibrium, the absence of perfect risk-sharing yields a non-degenerate cross-sectional distribution of income, consumption, and wealth, as well as individual mobility dynamics across the distribution, both of which resemble their data counterparts. As the bulk of macroeconomics of the last four decades, this class of models is also deeply rooted in the tradition of dynamic stochastic general equilibrium and rational expectations. More recently, economists have extended the analysis of the relation between distributions and monetary policy beyond the household sector to firms as well.[3]

The reason why central banks became engaged in this research program is that HANK models subvert some of the classic tenets of their representative-agent complete-market counterpart. A number of new policy lessons have emerged: (i) the transmission mechanism of monetary policy is no longer centered on intertemporal substitution, the dominant channel in RANK, but it revolves around equilibrium effects operating via shifts in labor income and asset prices; (ii) the cyclicality of income inequality and that of uninsurable labor income risk—both absent from RANK models—can substantially amplify the propagation of aggregate shocks; (iii) monetary policy leaves significant fiscal footprints because of the failure of Ricardian equivalence, a fixture of RANK; (iv) redistributive and social insurance policies

2. We refer the reader to Mankiw and Romer (1991) and Heathcote and others (2009) for an overviews of these two approaches.

3. See Ottonello and Winberry (2020).

are also aggregate stabilization policies and vice versa, i.e., the stark dichotomy between stabilizing business cycles and addressing imperfect insurance and inequality—which was an integral part of Samuelson's neoclassical synthesis—vanishes.[4] For a more detailed discussion of these new ramifications of HANK models, we refer the reader to Violante (2022).

The papers at the conference echoed all these messages. T.J. Sargent gave the keynote address at the conference. His paper, entitled **HAOK and HANK Models**, is a comparison—rich with informative historical references to the founders of modern macroeconomics—between the Heterogeneous-Agent Old Keynesian (HAOK) framework and the Heterogeneous-Agent New Keynesian (HANK) framework.

The old Keynesian paradigm is built on the idea that the macroeconomy can be, at times, in an equilibrium characterized by underemployment of capital and labor because of nominal rigidities. Other times, though, it operates efficiently under full employment, and markets alone are successful in setting prices correctly and allocating resources. In light of this observation, John Maynard Keynes advocated (what Sargent calls) "light-handed" monetary-fiscal interventions during downturns in order to restore full utilization and promote aggregate efficiency.

Paul Samuelson called this theory-policy pair a "neoclassical synthesis." At the heart of this view, Sargent argues, there is the implicit belief in the existence of well-functioning state-contingent transfers through markets, families or social safety nets that effectively insure households against adverse idiosyncratic shocks, such as job losses. As a result, macroeconomists could focus their attention on aggregate business cycles, without any reference to distributional issues. This perspective is what justified James Tobin's definition of macroeconomics as "a field that attains workable approximations by ignoring the effects of distributions on aggregates."

M. Friedman, R. Lucas, E.C. Prescott, and of course T.J. Sargent himself, together with other practitioners of twentieth-century macroeconomics, embraced this view. Most notably, they identified severe logical inconsistencies in early attempts to estimate Keynesian models through simultaneous equations systems. Their research program led to the rational expectations revolution and the paradigm that, seen through the lens of a structural model, time series are

4. For a more detailed discussion of these new ramifications of HANK models, we refer the reader to Violante (2022).

equilibrium stochastic processes. The class of dynamic stochastic general equilibrium models which emerged from that collective intellectual effort and which today pervades macroeconomics lays out environments where "a theory and an econometrics fit together consistently." These models still relied on the idea that the complete-market assumption offers a good approximation to actual economies, and thus the household sector of the economy could be collapsed into a representative agent. In his Presidential Address to the American Economic Association entitled "Macroeconomic Priorities", Bob Lucas wrote that "for individual behavior and welfare, of course, heterogeneity is everything. [...] for determining the behavior of aggregates, [...] household heterogeneity just does not matter very much" (see Lucas, 2003).

Then along came HANK models, which put heterogeneity and market incompleteness front and center in the study of business cycles. As a result of this additional complexity, solving and estimating these models requires new tools, many of which are showcased in the papers presented at the conference. According to Sargent, however, "the HANK revolution is not about tools but about substance. HANK research undermines the neoclassical synthesis." Sargent refers to the dichotomy between stabilization and redistribution implicit in that approach. In HANK economies, instead, traditional stabilization instruments such as countercyclical spending or interest rate cuts are necessarily redistributive and alter the amount of insurance against idiosyncratic shocks.[5] Similarly, traditional redistributive or social insurance instruments—such as tax reforms or expansion of unemployment insurance benefits—necessarily induce aggregate fluctuations because of the heterogeneity of marginal propensity to consume across the population.[6]

Sargent concludes his paper by airing the concern that this new theory-policy package could undermine traditional mandates for monetary policies and provide ammunition to constituencies that want to assign to central banks goals that involve redistribution and reallocation, with the risk of losing sight of price stability. While we share these

5. Bhandari and others (2021) study optimal monetary policy in HANK models and conclude that most of the gains accrue through improved consumption insurance, not higher aggregate efficiency.

6. There are, however, knife-edge cases where HANK models preserve this stark distinction, as articulated by Werning (2015).

concerns, we remain optimistic that policymakers will make good use of this new class of models without misinterpreting their implications. For example, HANK models can be beneficial to help choose between two policy interventions that attain, approximately, the same main objective (e.g., a certain trade-off between inflation and output) with different distributional consequences.

Motivated by the significant surge in energy prices in 2021 and the ensuing debate around the appropriate monetary and fiscal policy responses, the chapter by Adrien Auclert, Hugo Monnery, Matthew Rognlie, and Ludwig Straub (AMRS hereafter) titled **Managing an Energy Shock: Fiscal and Monetary Policy** examines the macroeconomic impact of energy price shocks in advanced, energy-importing economies. Their paper builds a HANK model of a small open economy that imports energy, by adding an energy good to the framework developed in Auclert and others (2021). This model allows the authors to explore how high energy prices may impact consumer demand, a recurrent concern voiced by policymakers. Such demand channel had remained largely unexplored either because existing work focused on the supply-side effects of energy price shocks, abstracting from nominal rigidities, or included sticky prices in RANK models where the demand channel is quantitatively trivial.

AMRS show how, under a realistic calibration of substitution elasticities and marginal propensities to consume (MPCs), energy price shocks can impact GDP via their effect on aggregate demand. Their main analytical result is that, when monetary policy keeps the real interest rate constant, the negative real-income effect (consumers demand less of all goods) dominates the substitution effect (households spend more on domestically produced goods) and aggregate output falls. In addition, aggregate dynamics do not display price-wage spirals because the recession caused by the shock pushes labor demand and wages down, thus offsetting workers' desire for higher wages linked to their decline in purchasing power. The paper also studies alternative monetary policy responses to the shock and uncovers interesting monetary and fiscal policy spillovers across countries. For example, while small individual countries acting alone would be unable to influence global prices, coordinated monetary tightening among energy-importing countries can reduce global energy demand, leading to lower energy prices and imported inflation. Turning to fiscal policies, AMRS argue that energy price subsidies can shield domestic consumers, but tend to have negative spillovers on other economies because they sustain the rise in world energy prices.

This paper offers an example of how models with realistic distributions of marginal propensity to consume can lead to a propagation mechanism of shocks that differs significantly from the one arising in representative-agent models.[7]

Measuring the Redistributive Effects of Monetary Policy: An Application to the Chilean Economy, by Emiliano Luttini, Ernesto Pastén, and Elisa Rubbo, explores the impact of monetary policy shocks across workers who are heterogeneous in their ex-ante (demographic) characteristics, consume different bundles, and work in sectors which differ in their capital intensity and their degree of nominal rigidity. Their multisector model, based on Rubbo (2023) and calibrated to the Chilean economy, shows that the response of employment and income of older, high-income men is almost 8 times larger than that of middle-aged, middle-income men. The reason for this unequal effect of monetary policy shocks can be almost exclusively traced back to the fact that the former group tends to work in industries with more severe nominal rigidities. This result hinges on the specificities of the Chilean economy and on the particular shock analyzed and, therefore, cannot be easily generalized. Taken at face value, though, it suggests that demand shocks might have a stronger impact on groups of workers that are, arguably, less liquidity-constrained and have lower MPCs. This particular distribution of exposure across households dampens the aggregate impact of the shock relative to a representative-agent model. A follow-up question for future research is whether the allocation of labor to sectors that differ in their exposure to demand shocks is an outcome of the optimal behavior of maximizing agents with different ability to self-insure.

The transmission mechanisms of monetary policy in environments with salient heterogeneity, this time in terms of financial intermediaries, is the topic of the chapter by Dean Corbae and Pablo D'Erasmo. In **The Bank Lending Channel Across Time and Space**, the authors set up an oligopoly model of heterogeneous banks with endogenous entry and exit to rationalize stylized facts about the U.S. banking sector after the Riegle-Neal Act, which permitted banks to cross state lines. The policy reform increased bank concentration at the national and state levels, but led to more geographic diversification of local shocks. The authors use the model to study the effects of this change in regulation on the bank lending channel of monetary policy.

7. Kaplan and Violante (2018) discuss various notions of equivalence between equilibrium outcomes in RANK and HANK models.

One important mechanism that generates substantial geographical heterogeneity in their model is that tighter monetary policy influences the equilibrium composition of the banking industry at the state level through the extensive margin, i.e., entry and exit. The reason is that large banks are less sensitive than small ones to a rise in the cost of external funds in the model, consistently with the microdata.[8]

The chapter by Sushant Acharya and coauthors entitled **Estimating HANK for Central Banks** provides a first assessment of the out-of-sample forecasting performance of HANK models, an operational issue that is at center stage for central banks. The authors use the HANK model of Bayer et al. (2024) which features the same types of shocks and frictions as the benchmark representative-agent new Keynesian (RANK) model of Smets and Wouters (2007). The paper makes a methodological contribution by explaining why and how the use of the Sequential Monte Carlo method can yield considerable efficiency gains when estimating HANK models. These gains are instrumental for the task of performing an out-of-sample assessment of these models, as one must estimate them multiple times.

Their main result is that no consistent improvement is found in the out-of-sample forecasting ability of HANK models. In fact, while for some series such as inflation, the forecasting ability is similar to that of standard RANK models commonly used by central banks, for other series, notably consumption growth, the performance is worse. This finding is surprising because the consumption block of the HANK model is much richer and more sophisticated than its RANK counterpart. The authors conjecture that a possible cause is that many parameters in their HANK model, namely all those affecting the model's steady state, continue to be calibrated ex ante.

They conclude that these results should motivate researchers to explore further ways to enhance the quantitative and out-of-sample properties of this class of models, which is still unsatisfactory. This is a worthy effort for a central bank which wants to fully understand the transmission mechanism of monetary policy, and its redistribution implications.

The standard view of macroeconomic dynamics—rooted in the monumental work of Burns and Mitchell (1946)—is that aggregate time series can be decomposed into a long-run component (the trend) and an orthogonal short-run component (the business cycle), which fluctuates around the trend. Quantitative DSGE models used for

8. See Kashyap and Stein (2000).

research and policy analysis fit into this description and, consistently with this view, routinely assume that transitory shocks have no long-term effects on aggregates.

A more nuanced view of business cycles is that of "macroeconomic hysteresis".[9] According to this interpretation, there is no longer a clear-cut separation between cycle and trend, and transitory shocks have very persistent, even permanent, effects on the level of economic activity. In their chapter entitled **From Micro to Macro Hysteresis: Long-Run Effects of Monetary Policy**, Felipe Alves and Giovanni L. Violante explore this alternative view by developing a HANK model built on the micro evidence that job losses lead to persistently lower individual earnings through a combination of skill decay and abandonment of the labor force. They show that these labor market micro-level sources of negative hysteresis give rise to macroeconomic hysteresis in response to transitory negative aggregate demand shocks, modeled as monetary policy innovations. In the model, the strength of these effects increases as one moves down the wage distribution: a decade after the shock, the scarring of labor earnings for workers in the lowest skill quartile is almost ten times as large as the average scarring effect. Hysteresis, thus, operates disproportionately through the labor market trajectories of low-wage workers. Despite the long shadow cast on output, the shock generates only short-lived movements in inflation, which quickly returns to its target. The reason for these dynamics is the decline in labor productivity and labor force participation, which jointly generate inflationary pressures that offset the long-run deflationary pull coming from the persistent decline in output.

Overall, the paper demonstrates that, thanks to their ability to richly represent heterogeneous exposure to aggregate shocks, HANK models are a natural laboratory to explore macroeconomic hysteresis that arises through the aggregation of microeconomic behavior in the labor market or, possibly, other markets.

There is much academic and policy interest centered on the question of how fiscal policy can be used to manage an economy that is stuck at the zero lower bound (ZLB). A number of classic results derived from RANK models prove equivalence, in terms of aggregate outcomes, between standard monetary policy, i.e., adjusting the nominal rate in response to demand shocks, and certain fiscal tools such as time-

9. See Cerra and others (2023) for a survey.

varying consumption subsidies.[10] These equivalence results are very useful to design stabilization policies when the ZLB binds, because they imply that these unconventional fiscal interventions can almost perfectly substitute for the lack of a monetary lever. In their chapter entitled **On the Optimal Use of Fiscal Stimulus Payments at the Zero Lower Bound**, Alisdair McKay and Christian K. Wolf revisit this question from the perspective of a HANK model. They study the optimal policy response for a government that wants to stabilize inflation and output but also dislikes consumption inequality in excess of its steady-state level. Their key conclusion is that, for canonical ZLB-type shocks—like a tightening in borrowing constraints or a distributional shock concentrated on low-income households—the best alternative to classical unconstrained monetary policy is not consumption subsidies, but uniform transfer stimulus payments. The reason is that, beyond perfectly stabilizing aggregate output and inflation, they boost consumption of low-income households, directly counteracting the distributional incidence of the original business-cycle shock. As a result, stimulus payments do not just substitute for conventional monetary policy—they strictly improve upon it.

This paper is a stark example of how HANK models can be useful to policymakers in choosing among alternative policies that achieve very similar aggregate outcomes, but yield different distributional implications.

Another example of this logic can be found in **The Role of Progressivity on the Economic Impact of Fiscal Transfers: a HANK for Chile**, by Benjamin García, Mario Giarda, and Carlos Lizama. The authors start by documenting a strong non-Ricardian response of the Chilean economy to fiscal transfers. In addition, they find that more progressive fiscal transfers display significantly larger effects on consumption than less progressive ones. Motivated by these empirical results, they set up a HANK model with search and matching frictions and calibrate it to key moments of the Chilean economy, such as the fraction of hand-to-mouth households from household surveys and income dynamics from administrative data. The model is able to reproduce the main finding that fiscal transfers geared towards households with higher MPCs have a larger macroeconomic impact. Furthermore, they show that this impact is amplified if transfers are financed through debt instead of taxes. The authors conclude by speculating that the right combination of expansionary fiscal and

10. See Correia and others (2013).

contractionary monetary policy could entail significant redistribution without displaying large adverse aggregate effects.

Overall, the Central Bank of Chile conference showcased a rapidly evolving and vibrant field that is challenging some acquired wisdom, while remaining well anchored to the successful research program of dynamic stochastic general equilibrium models in macroeconomics. Its success will depend on two factors. First, the extent to which practitioners of these models will be able to make a convincing and robust case that a two-way feedback between inequality and the macroeconomy exists and is quantitatively important. Second, the development of better computational and econometric tools to solve globally and estimate stochastic models. Recent advances based on neural nets appear to offer a promising avenue.

References

Auclert, A., M. Rognlie, M. Souchier, and L. Straub. 2021. "Exchange Rates and Monetary Policy with Heterogeneous Agents: Sizing Up the Real Income Channel." Technical Report, National Bureau of Economic Research.

Bayer, C., B. Born, and R. Luetticke. 2024. "Shocks, Frictions, and Inequality in U.S. Business Cycles." *American Economic Review* 114(5): 1211–47.

Bhandari, A., D. Evans, M. Golosov, and T.J. Sargent. 2021. "Inequality, Business Cycles, and Monetary-Fiscal Policy." *Econometrica* 89(6): 2559–99.

Burns, A.F. and W.C. Mitchell. 1946. "Measuring Business Cycles." National Bureau of Economic Research.

Cerra, V., A. Fatás, and S.C. Saxena. 2023. "Hysteresis and Business Cycles." *Journal of Economic Literature* 61(1): 181–225.

Correia, I., E. Farhi, J.P. Nicolini, and P. Teles. 2013. "Unconventional Fiscal Policy at the Zero Bound." *American Economic Review* 103(4): 1172–211.

Heathcote, J., K. Storesletten, and G.L. Violante. 2009. "Quantitative Macroeconomics with Heterogeneous Households." *Annual Review of Economics* 1(1): 319–54.

Kaplan, G. and G.L. Violante. 2018. "Microeconomic Heterogeneity and Macroeconomic Shocks." *Journal of Economic Perspectives* 32(3): 167–94.

Kashyap, A.K. and J.C. Stein. 2000. "What Do a Million Observations on Banks Say about the Transmission of Monetary Policy?" *American Economic Review* 90(3): 407–28.

Lucas, R.E. 2003. "Macroeconomic Priorities." *American Economic Review* 93(1): 1–14.

Mankiw, N.G. and D. Romer. 1991. New Keynesian Economics: Coordination Failures and Real Rigidities." Boston, MA: MIT Press.

Ottonello, P. and T. Winberry. 2020), "Financial Heterogeneity and the Investment Channel of Monetary Policy." *Econometrica* 88(6): 2473–502.

Rubbo, E. 2023. "Monetary Non-Neutrality in the Cross-Section." Technical Report.

Smets, F. and R Wouters. 2007. "Shocks and Frictions in U.S. Business Cycles: A Bayesian DSGE Approach." *American Economic Review* 97(3): 586–606.

Violante, G.L. 2022. "What Have We Learned from HANK Models, thus Far?" In *Beyond the Pandemic: the Future of Monetary Policy*. ECB Forum on Central Banking 2021, European Central Bank, Frankfurt, Germany.

Werning, I. 2015. "Incomplete Markets and Aggregate Demand." Technical Report, National Bureau of Economic Research.