

Macrofinance Dynamics, Heterogeneity,  
and Policy Design

Dissertation zur Erlangung des Grades eines  
Doktors der Wirtschaftswissenschaften  
der Universität Bielefeld

vorgelegt von  
Gregor Böhl

Bielefeld, Juni 2017

<b>Erstgutachter:</b>	Prof. Dr. Herbert Dawid	Universität Bielefeld
	Prof. Dr. Cars Hommes	Universität von Amsterdam
<b>Zweitgutachter: :</b>	Prof. Dr. Cees Diks	Universität von Amsterdam
<b>Sonstige Mitglieder:</b>	Prof. Dr. Christiane Clemens	Universität Bielefeld
	Prof. Dr. Alfred Greiner	Universität Bielefeld
	Dr. Christian Stoltenberg	Universität von Amsterdam
	Prof. Dr. Jan Tuinstra	Universität von Amsterdam
	Dr. Florian Wagener	Universität von Amsterdam

This thesis has been written within the European Doctorate in Economics - ErasmusMundus (EDEEM), with the purpose to obtain a doctoral degree in economics at the University of Amsterdam and at Bielefeld University.

Wirtschaftswissenschaftliche Fakultät, Universität Bielefeld

Diese Arbeit wurde auf alterungsbeständigem Papier gemäß ISO 9706 gedruckt.

# Abstract

The present dissertation contains four chapters: the first chapter is an introduction to motivate this dissertation and interlink the different chapters with each other. The second chapter is associated with the role of monetary policy when stock markets are driven by boundedly speculative traders. The third chapter explores the market outcome if fully rational and boundedly rational agents interact in a financial market. The last chapter is concerned with the dynamics of the distribution of wealth given capital gains taxation.

In the first chapter, the motivation of this thesis is briefly summarized. In particular, both the 2007 – 2010 financial crisis and the publication of the Piketty data had a major impact on economic science and called out for a shift in paradigms. For one, this had led to the inclusion of a financial sector and respective frictions into dynamic macro models, but also to a more careful treatment of heterogeneity. Further, this has led to methodological advances, especially with regard to techniques that allow the analysis of nonlinear Rational Expectations system and tools that enable properly to model heterogeneity.

The second chapter studies the macroeconomic consequences when – potentially speculative – stock prices affect macroeconomic aggregates, and whether monetary policy can mitigate potential spillovers from financial markets. I augment a model with financial constraints on working capital with stock markets, where excess volatility of these markets is endogenously amplified through behaviorally motivated financial speculation. The presence of credit constraints links asset returns to optimal leverage and the price level. Then, standard monetary policy rules can induce a dynamic feedback loop that actually amplifies stock price volatility. This model is estimated to match key moments of empirical European macroeconomic data. The endogenous process of financial market speculation and the feedback from asset prices to the price level are key features to replicate and explain these moments. Numerical analysis suggests that central banks can offset the impact of speculation on either output or inflation by carefully targeting asset prices, but not on both, and can furthermore dampen excess volatility of stock prices. However, the scope of such policy to stabilize economic activity is limited due to its undesirable response to real economic shocks.

Chapter three analyses the interaction of perfectly rational agents in a market with coexisting boundedly rational traders. Whether an individual agent is perfectly rational or boundedly rational is determined endogenously depending on each types market performance. Perfect rationality implies full knowledge of the model including the non-linear switching process itself. I make use of iterative numerical methods to find a recursive solution of the highly nonlinear system, which then only depends on the state space of the original model, and show furthermore that this solution is not necessarily bounded. Depending on the parameterization, agents' interaction can trigger complicated endogenous fluctuations that are well captured by our solution algorithm. In conclusion, in a financial market setup rational agents might adapt sentiment beliefs and so fail to mitigate speculative behavior, and boundedly rational agents are not necessarily driven out of the market. Although the presence of fully rational agents tends to have stabilizing effects it may also induce further endogenous fluctuations.

In the forth and last chapter, using a parsimonious economic model we are able to explain up to 100 years of the available data on the dynamics of top-wealth shares for the US, the UK, and France both in levels and transitions. We build a microfounded model of heterogeneous agents where additionally to stochastic returns on investment individuals disagree marginally on their expectations of future returns and thus hold different portfolios. Closed-form solutions confirm that without government intervention this process has a variance that explodes in time implying no finite inequality. We then introduce a tax on capital gains and provide evidence that this converges to a double Pareto distribution for which the degree of wealth inequality is decreasing in the tax rate. We provide numerical simulations for the calibrated model as well as analytical representations for transitions and cross-sectional distributions, and discuss its ability to match the measured wealth inequality for several countries. The heterogeneous development in the different countries and across time can be traced back very precisely to different tax regimes.



## Acknowledgements

At the time I started my PhD I thought I was smart. Not sure about that retrospectively. Thinking about it, I always thought I was smart, and yet time and time again I got convinced I was not. In fact, I still think that I am smart, but finally I learned to mistrust this idea. It took me a PhD to realize that. These last 4 years have been an extraordinary period, with so many moments of joyful thinking and of gaining deeper understanding. Equally deeply, I am indebted to those people who made this experience possible, and to those who companioned me along my path.

I want to sincerely thank both of my supervisors Herbert Dawid and Cars Hommes. Herbert provided me with all the support (and more) necessary so that I did not have to worry for a second about anything else than how to put my ideas to work. I believe that there is no better match than a PhD student who wants to figure things out on his own and a supervisor who is able to put his trust on him. I am very happy about that, given how many graduate students are suffering from too-much imposing supervisors. Cars adopted me half-way through my PhD and gave me the indispensable chance to come to Amsterdam for new input, experiences, and fruitful contacts. From both I learned not only about methodology and scientific thinking but they also thought me that science sometimes might be more than just having good ideas and writing them down. Both tried to teach me how to present my thoughts clearly and comprehensively, and the fact that this *proefschrift* might prove otherwise can only be attributed to my own impatience.

Furthermore I am very grateful to Christian Stoltenberg for his valuable lessons on how to think like a macroeconomist and his faith in me. Besides this I also much owe him for his support in the job market. Thanks to his effort jointly with my supervisors that I can continue along this path in an excellent environment. Nikos Georgantzis has, since his supervision of my Master's thesis in Spain, always promoted me for no good reason and it was him who actually pointed me to Herbert Dawid in the very beginning. For some reason, he always believed in me and was never hesitant to provide support. My co-promotor Cees Diks was a source of motivation, technical enlightenment and amusing geeky chats, and it was of great help that he would take the time for a discussion or give a hint whenever necessary.

A special thanks goes to Alex Clymo, from whom I have learned so much in discussions over uncountably many refreshing beverages and who was never shy of an advice. I hope our friendship will prevail, not the least to spread the truth in our songs that was yet untold. I thank Urs Schröder, a very gifted man in so many ways, for selflessly setting the stage to live a lifestyle that was extraordinary (and everybody else that lectured me on "all about rust"). I am very happy that Thomas Fischer joined me as a coauthor for my last chapter and I am glad that this collaboration turned out fruitful.

Along the way I have met so many people that I owe a big thank you and sometimes also an apology. In Bielefeld I was happy to meet Sander van der Hoog who inspired me in my early stage and supported me unselfishly in difficult situations. Philipp Harting was a friend, an instance and an indispensable daily routine which I otherwise would have missed a lot. I had the delight to stare at the black square with Jan Lohage, who was not only refreshing but also leveled-up my geekyness. Thanks to Christoph, Johannes, Rike, Sarah and Steffi and everybody else who stressed the word *köstlich* in the last years for retrospectively making my relocation to the province a good idea. A special thanks goes to those who sacrificed their souls, brains and probably also body parts to help me translating the summary of this thesis. Nobody can imagine the pain they went through. All remaining errors are mine (and I'll keep them).

Karo was always there for me and I would immediately go to the end of the world with you. Myrna was a calm and encouraging companion at travels, and probably she is the kindest person that walks on this planet. Stephanie shared with me all the crucial knowledge on unicorns (which is a lot). May the holy diver stand by your side for evermore. I was also happy to meet Sara again in Amsterdam. Without her, the start would not have been easy. Thanks to Daniela for convincing me that there are parts of the Netherlands which are actually quite nice (and sorry for being too slow). Moutaz was the best office mate anyone could hope for and meeting him was not only a cultural enrichment. Oana and me still have a lot of songs left to sing. In Amsterdam I further had the pleasure of having Egle, Gavin, Marcello and Swapnil around, who certainly made life easier.

It goes without saying that without my new and old friends Charlotte, Cia, Fabian, Inga, Julian, Martin, Nils, Karen, Kiki and the Vocals these would have been boring years.

Nicht zuletzt danke ich meinen Eltern Monika und Klaus für ihre Unterstützung und ihren Glauben in mich. Ich hatte immer die Freiheit, meine Wege selbst wählen zu dürfen und dadurch kam nie das Gefühl auf etwas verpasst zu haben.

## Coauthors

Chapter 4, entitled “*Capital Taxation, Investment and the Dynamics of the Wealth Distribution*” is joint work with Dr. Thomas Fischer from Lund University. It is based on a simpler idea which we jointly elaborated upon. We extended the model together and Thomas added his excellent knowledge of the literature, data sources and analytical techniques. Apart from the idea and concept, I am responsible for the majority of the analytical and programming part. The writing part was approximately split into half.

# Contents

Abstract . . . . .	iii
Acknowledgements . . . . .	v
List of Tables . . . . .	xii
List of Figures . . . . .	xiii
<b>1 Introduction</b>	<b>1</b>
<b>2 Monetary Policy and Speculative Stock Markets</b>	<b>5</b>
2.1 Introduction . . . . .	5
2.2 Model . . . . .	12
2.2.1 Households . . . . .	12
2.2.2 Firms . . . . .	13
2.2.3 Financial Intermediation . . . . .	18
2.2.4 Central Bank and Government . . . . .	19
2.3 General Equilibrium and Estimation . . . . .	20
2.4 Endogenous Fluctuations in Asset Prices . . . . .	24
2.4.1 Heterogeneous beliefs . . . . .	29
2.4.2 Deterministic Simulations . . . . .	32
2.4.3 Policy and Stochastic Simulations . . . . .	35
2.5 Conclusion . . . . .	39

<b>Appendix A</b>	<b>41</b>
A.1 Entrepreneurs' optimization problem . . . . .	41
A.2 Bayesian estimation . . . . .	43
A.3 Solving for the rational expectations equilibrium . . . . .	45
A.4 Eigenvalues of the Rational Expectations System . . . . .	47
A.5 Fitting the model to match the statistical moments from Table 1 . . . . .	47
<b>3 On the Evolutionary Fitness of Rationality</b>	<b>49</b>
3.1 Introduction . . . . .	49
3.2 Model . . . . .	52
3.3 Numerical solution . . . . .	55
3.4 Results . . . . .	58
3.4.1 Endogenous dynamics . . . . .	59
3.4.2 The consequences of rationality . . . . .	61
3.4.3 Comparing rational and fundamentalist traders . . . . .	62
3.4.4 The two-trader type model . . . . .	62
3.4.5 Risk adjusted fitness measure . . . . .	65
3.5 Conclusion . . . . .	67
<b>Appendix B</b>	<b>69</b>
B.1 Intuition behind Proposition 1 . . . . .	69
B.2 A note on convergence . . . . .	71
<b>4 Capital Taxation and Investment: Matching 100 Years of Wealth Inequality Dynamics</b>	<b>73</b>
4.1 Introduction . . . . .	73
4.2 Literature . . . . .	76
4.3 Model . . . . .	80
4.4 Analytical Results . . . . .	85
4.5 An Empirical Application . . . . .	89

4.5.1	The United States . . . . .	94
4.5.2	The United Kingdom . . . . .	96
4.5.3	Sweden . . . . .	97
4.5.4	France . . . . .	99
4.6	Validation and Projections 2040 . . . . .	100
4.7	Conclusion . . . . .	103
<b>Appendix C</b>		<b>105</b>
C.1	Proofs of the Propositions . . . . .	105
C.1.1	Proof of Proposition 2 . . . . .	105
C.2	Robustness with Respect to $\gamma$ . . . . .	110
C.3	Out-of-sample tests for the UK and France . . . . .	111
C.4	Additional Notes on the Stationary Distribution . . . . .	112
<b>Bibliography</b>		<b>115</b>

# List of Tables

2.1	Standard deviations and cross correlations of inflation, output and real stock prices, Core-Europe from 1976 to 2014 (quarterly). A detailed description of the data can be found in Section 2.3. . . . .	9
2.2	Parameter estimates of the RE model using a grid-based minimizing procedure. . .	23
2.3	Standard deviations & cross correlations of the estimated Rational Expectation model. The model is unable to correctly capture the covariance between stock prices and inflation. . . . .	24
2.4	Parameter estimates for the model with endogenous dynamics in stock prices, based on a grid-based minimization procedure. . . . .	32
2.5	Standard deviations and cross correlation matrix for simulations (10.000 iterations) of the model with endogenous dynamics in stock prices. . . . .	32
A.1	Priors and parameter estimates of the Bayesian estimation . . . . .	44
3.1	Benchmark parametrisation . . . . .	58
4.1	Parameter estimates used for closed-form low-of-motion and simulations. For the UK the analytic solution underestimates $\gamma$ , which is why for the simulations $\gamma = 0.082$ is used. . . . .	93



# List of Figures

2.1	General equilibrium effect of a 1% change in asset price expectations as a function of central bank policy parameters. Responses of output, inflation and asset prices in deviations from steady state. . . . .	28
2.2	Long-run deterministic dynamics of $y_t$ (blue/light) and $\pi_t$ (green/dark) with respect to the behavioral bias $\alpha$ . $\gamma = 1.29$ , all other parameters as in Table 2.4. . . . .	33
2.3	Deterministic dynamics for $\phi_s$ . $\alpha = 1.31$ and $\gamma = 1$ , all other parameters as in Table 2.4. . . . .	35
2.4	Standard deviations of stochastic simulations for given the asset price targeting policy parameter ( $\phi_s$ ) and the model with endogenous dynamics in stock prices. The economy is driven by exogenous shocks, but the nonlinear process in the stock market leads to endogenous amplification. . . . .	38
A.1	Eigenvalues of the underlying model ( $\mathbf{N}^{-1}$ ) as a function of monetary policy parameters and the elasticity of the external finance premium. . . . .	47
A.2	Norm of the Matrix of squared errors. Here as a function of the marginal cost elasticity to stock prices . . . . .	48
3.1	Bifurcations w.r.t. $\delta$ . In blue (green) and on the right axes the dynamics of the fraction of rational (sentiment) traders. . . . .	59
3.2	Bifurcations w.r.t. $\beta$ . In blue and on the right axes the dynamics of the fraction of rational traders. . . . .	59

3.3	Bifurcations w.r.t. $\kappa$ . $\beta = 1.05$ . Blue and green are the fractions of rational and positive biased agents. . . . .	61
3.4	Bifurcations w.r.t. $\kappa$ . $\beta = 2$ . In blue dynamics of the fractions of rational agents. . .	61
3.5	Bifurcations w.r.t. $\delta$ . In cyan the same simulation with fundamentalists instead of rational agents. . . . .	63
3.6	Heatmap of $g_{75}(x_{t-1}, x_{t-2}, 0)$ for $\gamma = 0.2$ (l) and $\gamma = 0.25$ (r). For the lower gamma, the zero steady state is still locally stable. For $\gamma = 0.25$ trajectories diverge slowly outwards. Note the different scales. . . . .	63
3.7	Heatmap of $g_{75}(x_{t-1}, x_{t-2}, 0)$ for $\gamma = 0.32$ (l) and $\gamma = 0.4$ (r). Trajectories lead from the center to the periphery. This effect amplifies with higher $\gamma$ . Note the different scales. . . . .	64
3.8	Bifurcation diagrams for $\delta$ (left), $\beta$ (center) and $\kappa$ (right). For the first two parameters the steady state increases with the parameter but falls back after a certain threshold. The last Figure depicts the case where $\delta = 5$ . The blue line shows the fraction of rational agents. . . . .	66
B.1	Illustration of the numerical function $g$ for $\delta = 1.1$ (ul), $\delta = 2.5$ (ur), $\delta = 4$ (ll) and $\delta = 7$ (lr). When $\delta$ increases the function becomes steeper and at the diagonal a small change in $x_{t-1}, x_{t-2}$ leads to a larger change in $x_t = g(x_{t-1}, x_{t-2})$ . . . . .	71
4.1	Top wealth shares. Data source: <code>wid.world</code> and Lundberg and Waldenström (2017) for Sweden. . . . .	77
4.2	Trajectories using the law-of-motion in closed form (Equation 4.4.11) and the parameters from Table 4.1. Each time series is initialized with the top-share from the empirical data and then uses the available series of (capital) income taxes as the single input. . . . .	94

4.3	Closed form LOM $s_t^{0.01}(\tau_t)$ (solid line) and closed form stationary share $s^{0.01}(\tau_t, \infty)$ (dotted line) taking tax series and the calibration for the USA together with the empirical evidence (dashed line) and the 95% interval of the respective simulations.	95
4.4	Simulations starting from the wealth distribution in 1954 for the USA, using only the respective time series of US capital gain taxes as input ( $\gamma = 0.3523$ ).	95
4.5	Closed form LOM $s_t^{0.01}(\tau_t)$ (solid line) and closed form stationary share $s^{0.01}(\tau_t, \infty)$ (dotted line) taking tax series and the calibration for the UK together with the empirical evidence (dashed line) and the 95% interval of the respective simulations.	97
4.6	Simulations starting from the wealth distribution in 1913 for the UK, using only the respective time series of UK top income taxes as input. $\gamma = 0.082$ and $\eta$ as in Table 4.1.	97
4.7	Closed form LOM $s_t^{0.01}(\tau_t)$ (solid line) and closed form stationary share $s^{0.01}(\tau_t, \infty)$ (dotted line) taking tax series and the calibration for Sweden together with the empirical evidence (dashed line) and the 95% interval of the respective simulations.	98
4.8	Simulations starting from the wealth distribution in 1908 for Sweden, using only the respective time series of Swedish top income taxes as input. Parameters as in Table 4.1.	98
4.9	Closed form LOM $s_t^{0.01}(\tau_t)$ (solid line) and closed form stationary share $s^{0.01}(\tau_t, \infty)$ (dotted line) taking tax series and the calibration for France together with the empirical evidence (dashed line) and the 95% interval of the respective simulations.	100
4.10	Simulations starting from the wealth distribution in 1917 for France, using only the respective time series of French top income taxes as input. Parameters as in Table 4.1.	100
4.11	Out-of-sample tests for the US (left) and Sweden (right). The fat bars indicate the sample period.	102

4.12	At the crossroads: projections for different tax regimes. For the US (left) the top dashed line projects the current tax rate, the second line shows that a rate of at least 16.6% would be necessary to keep inequality at the present level. A rate of 20% would partly reverse the trend. The bottom line shows a 25% rate. For Sweden a rate of 56.6% is required to keep inequality at the present level. . . . .	103
C.1	Variations of $\gamma$ in the closed-form solutions for the USA (left) and the UK (right).	110
C.2	Variations of $\gamma$ in the closed-form solutions for Sweden (left) and France (right).	111
C.3	Out-of-sample predictions for the UK (left) and France (right). Fat bars indicate the respective sample period. . . . .	112
C.4	Log histogram of simulation using US calibration after a very long horizon of 50,000 periods. $\gamma = 0.3523$ , $\tau = 25\%$ and $N = 100,000$ . . . . .	113
C.5	Right tail of the distribution for the last period for exemplary simulation using the US calibration of $\gamma = 0.3523$ , $\bar{\omega}_{US}^{-1} = 0$ and $\theta_t = 0.25$ at all $t$ . The dashed line represents a Pareto-fit with an estimated coefficient of 1.459. . . . .	113

# Chapter 1

## Introduction

As many of the doctoral students in economics that finish their theses in these years, my early economic studies have been branded by those events that surfaced in the second half of the last decade, but are the aftermath of processes that started earlier. The collapse of the real estate markets, the cascading effects of banking networks, the sovereign debt crises and the crises of the European currency union form an integrative part of the collective memory of my generation of economists. Devastating in their consequences for Europe and the World, those events presented great challenges for a young economist, but also temptations. Since the "old" theories did not seem to work anymore there was a vacuum in theory feasible to give policy advice, and we had the great pleasure to be forced to think outside the box in developing new theory and methodology.

Although already introduced and discussed into the economic literature earlier, economists were led by the urge to provide meaningful tools and started to put more effort into integrating financial frictions into macroeconomic models. In a broad sense, the second chapter of this thesis contributes to this branch of research. After 2008 it was not negligible that under certain circumstances the financial structure of firms and household plays a crucial role for the macroeconomy as well. While most recent publications that include financial frictions, starting prominently with Bernanke et al. (1999), focus on the role of external finance and collateral, I focus on the effects of fluctuations in stock prices which are relevant for internal finance and raising funds at the capital market.

By construction, stock prices can only make a difference in a macroeconomic model if we abstain from the assumption that they perfectly reflect future profits. This is in particular, one of the primary causes of the financial crisis was the over valuation of financial assets that were traded among banks, in combination with the failure of rating agencies to correctly account for the risk of these assets. The long history of financial bubbles and an immanently growing literature on bounded rationality suggest that the assumption of rational evaluation of assets might be controvertible. Already Keynes (1936) stressed the nature of financial markets to resemble a game of *beauty contests*, a notion that can be formalized as a *positive feedback loops*. I argue that such feedback loop in the stock market enhances the probability of boundedly rational behavior, and might have strong consequences for the macroeconomy as such. To study potential policy responses I allow the monetary authority to not only adjust the policy rate with regard to changes inflation but also to stock prices in order to mitigate this feedback. The results of this policy experiment are presented in Chapter 2.

A natural candidate when explaining bubbles and sudden crashes is bounded rationality. I am treading the interaction of boundedly rational vs rational agents in Chapter 3 of this thesis. The intuition behind this interaction problem is quite straightforward. Given that it is known that some of the participants of a financial market are biased concerning future returns, how would a rational trading decision look like? Rationally, if some agents are optimistic, the price would be expected to increase and hence a rational trader is incentivized to raise demand. However, the resulting equilibrium effect would actually confirm the optimistic belief – a vicious circle emerges. But given that argument, should not prices then always be unstable, because already a single boundedly rational agent in an otherwise rational market would induce such upward spiral? I am studying this question of stability and dynamics in a setup not only with biased agents, but in a framework of both, bounded and complete rationality, and endogenous switching. Since the behavior of boundedly rational agents affects the decisions made by rational agents, the beliefs of the boundedly rational agents remain self-fulfilling and they remain in the market. This result is not only relevant for asset pricing theory, but also for macroeconomics and macrofinance.

During my PhD studies another event occurred that likewise had an impact on economic science and my own work: Piketty (2014) published his large collection of data on income and wealth inequality. Almost neglected before, questions of income and wealth distributions returned to the focus of economic debates. Apart from the discussion on impact of distributional measures on macroeconomic aggregates, this collection now allows to access large amounts of data and time series on the dynamics of inequality in the last century. Chapter 4 of this thesis is devoted to that topic where I, together with my coauthor Thomas Fischer, develop a parsimonious model of risky investment that dynamically maps the series of capital income taxes into wealth inequality. The outcome of this model can then be compared, related and estimated to and with the Piketty data. Although much theory has been produced to explain wealth inequality, we suggest that the time series on taxes alone is to a large extent sufficient to explain dynamics of the wealth distribution, both in terms of levels and transitions.

These shifts in macroeconomic theory, further amplified by the rapid increase in computational power in the last 10 – 20 years, went along with respective shifts of methodology. This has in particular led to boost the use of computational methods. For one, to find the Rational Expectations equilibrium of a nonlinear system or a system involving real heterogeneity is a nontrivial task where a closed-form expression of the law of motion is normally not available. Further, to model a large population of agents within a feasible requires advanced skills such as proper calibration and efficient programming.

In this thesis the relevance of heterogeneity is emphasized in particular, both among agents stocks but also in their decision rules. In chapter 2 I make use of more traditional methods to solve linearized rational expectations models, but combine those with the nonlinear dynamics induced by heterogeneity in expectations. The third chapter applies recursive iterative methods to find solution of the law of motion of a speculative market that embeds fully rational as well as boundedly rational agents. In the fourth chapter we simulate large (50.000+) populations of agents over a long horizon to capture the dynamics of inequality, in addition to analytical methods adapted from physics science.

Each of the chapters of this thesis can be read independently, providing an individual introduction, conclusion and appendix. To avoid redundancy the bibliography is summarized at the end of this thesis.



# Chapter 2

## Monetary Policy and Speculative Stock Markets

### 2.1 Introduction

*“I consider asset price bubbles and monetary policy to be one of the most challenging issues facing a modern central bank at the beginning of the 21st century.”*

— Jean-Claude Trichet, 2005

Observers of the great recession have argued that when facing a liquidity trap, expansionary monetary policy will, instead of fostering steady growth, tend to fuel financial markets. These in turn might further destabilize the economy and comprise hazard. Some economists (Borio and Lowe, 2002; Cecchetti, 2000) have suggested to let monetary policy target asset prices to prevent the aforementioned spiral of bubbles, instability and unconventional monetary policy. When suspecting such a feedback between financial markets and real aggregates, two postulates are implied. The first postulate is the existence of a mutual link between asset prices and real activity, i.e. that causality might run in both directions. Recent economic literature stresses the relevance of firms’ financial structure. This suggests to also study the macroeconomic impact of equity, and equity prices. Yet there is limited insight on how and when stock prices impact real

activity. The second postulate is that asset prices do not only reflect a discounted fundamental value but embed distortions themselves, possibly biased by speculation. Such destabilizing speculative process implies that traders are not fully rational and that the existence of potential positive profits through speculation makes financial markets more prone to instability than real markets.

This work tries to answer the question whether causality does not only run from real activity to stock markets, but also in the opposite direction. Starting with the microfoundations of such linkage, I use a parsimonious macroeconomic model to replicate statistical key-moments of the data. As it turns out, both the linkage as well as a speculative process are necessary to sufficiently match the data. Having thus motivated a positive role for policy, I ask whether a Taylor-type interest rate rule that also targets asset prices is able to mitigate the impact of speculation on real activity and to reduce excess volatility on stock markets.

The first contribution of this work is hence to identify the potential channels of how stock prices could feedback into macroeconomic activity. The first channel transmits through the *external finance premium* introduced by Bernanke et al. (1999) (BGG). When firms' financial structure is relevant and borrowing constraints depend on the quality of equity, an increase in equity prices lowers costs for external funding. Secondly, since stock market prices reflect future expected returns, potential holders will only purchase shares of those firms that reflect the highest expected return per invested unit. If certain shares are valued below the market price, holders are incentivised to liquidate the firm and reinvest the return in a firm with higher return to equity. If firms are ex ante identical, all firms must provide the same return on equity.<sup>1</sup> Further below I show that both concepts result in the same mathematical model.

I formulate a DSGE model of a monetary production economy and extend it by a sector of financial intermediaries. Firms lever their profits by borrowing from the financial intermediaries and pledge their equity as collateral, while borrowing conditions and finance costs depend on the quantity of collateral offered. The external finance premium thus depends on firms net worth. I

---

<sup>1</sup> A further potential channel is a wealth effect that works through aggregate demand: increasing stock prices raise the nominal value of assets held by households, and amplifies consumer demand. Unfortunately such effect is ruled out in a representative agent framework where seller and buyer are identical and changes in asset prices level out to zero in aggregate. Since that, an increase in asset prices falls short to increase households' real spending opportunities.

assume that firms issue equity shares and can choose their net worth by deciding over dividends paid to shareholders. It is furthermore assumed that firms can also raise capital from shareholders. If firms seek to maximize the dividends per share, I show that under reasonable general assumptions asset prices are linked to the profit rates and are competitive among firms. This connects asset prices and return on equity, whereas optimality requires return on equity to equal the external finance premium. Return on equity in turn depends on price-setting which, in aggregate, determines the consumer price level. Thus, if firms maximize expected returns on equity, a link between stock prices and aggregate economic activity is probable.

The second, methodological contribution concerns the interaction of speculative and rational agents. Financial markets work fundamentally different than commodity markets. While commodities are a means to an end, notably consumption or production, people hold assets because of their expected return in the future. Hence, their value is directly expressible in monetary terms. Focussing on the resale value of a financial asset reveals that beliefs about prices are to a certain extent self-fulfilling. Furthermore, incentive schemes differ notably between financial assets and real goods. To illustrate, let us imagine a firm that has anticipated the current price correctly in the previous period. Then profits are maximal since the optimal production volume is chosen with respect to the firm's cost function. But if prices were overestimated, the firm will be unable to sell the produced stock profitably and incurs a relative loss. Asset markets work differently. If a positive price change has been overestimated by a trader, he will still realize a higher profit than a second trader that expected the price change correctly. Unlike commodity markets, traders in asset markets can benefit from overoptimistic forecasts at least in the short run. This can lead to herding behaviour: instead of focussing on the underlying fundamental, it can be *behaviorally rational* (Hommes, 2013) for traders to follow the majority in their beliefs. Such a mechanism is not well captured by rational expectations. To incorporate these anomalies and to allow for realistic asset price dynamics I make a distinction between expectations on *real economic aggregates* like output and inflation and expectations of financial market prices. While expectations on the real side of the economy are modeled to be perfectly rational, financial traders form boundedly rational

expectations that embed speculative dynamics. This enables to explicitly study the feedback between stock prices and real aggregates to derive implications for monetary policy. Restricting the bounded rationality to a small part of the model addresses the prominent critique of *the wilderness of bounded rationality* (Sims, 1980) while retaining a strong forward-looking component in the framework. It enable me furthermore to make use of nonlinear dynamic theory to identify both, qualitative and quantitative changes in dynamics depending on the central banks policy rule.

My third and central contribution is to study the role of monetary policy to stabilize the economy via asset price targeting. The fundamental idea behind this concept is that the interest rate impacts on asset prices through the discount factor, i.e. a higher interest rate leads to a devaluation of assets. This then keeps firms from being levered too much in boom phases, and increases the buffer in bad times. In fact, I show that conventional monetary policy can amplify fluctuations in stock prices if they affect the firms financial costs. An increase in stock prices decreases firms marginal costs of finance and hence, if prices are linked to marginal costs, can induce a decrease in consumer prices. When monetary policy then reacts with a decrease in interest rates, this increases stock prices even further.

## **Related Literature**

Winkler (2014) uses BGG-type frictions to combine asset prices and real activity, and empathizes the role of learning in an otherwise rational model. This enables to reproduce excess volatility of asset prices as well as a relatively high standard deviation of stock prices. As in my model, under rational expectations a monetary policy that targets asset prices induces a welfare-loss, while it is argued that under learning, carefully targeting asset prices might lead to a welfare improvement. It is however noteworthy that Winkler (2014) and others are using US data while I am focussing on the case for the Euro area. Miao et al. (2012) and Miao et al. (2016) build a Bayesian model with rational stock price bubbles which also affect the economy through endogenous borrowing constraints. Similar to my model, the feedback between asset prices and asset price expectations plays a key role on the formation of a stock price bubble. Compared to this literature, the emphasis

of my analysis lies on two effects: first, under which conditions a monetary policy that targets asset prices could stabilize financial markets and as such mitigate the intensity of spillovers? And secondly, whether and how it is possible to unlink economic aggregates from asset prices such that excess volatility would not matter for economic welfare. I show that this question is closely interrelated to this feedback loop, because asset price targeting can effectively mitigate the direct feedback of expectations on current prices, which is important independently of whether a bubble is due to bounded rationality or other mechanism.

A different branch of the literature most prominently represented by Galí (2013) uses the concept of rational asset price bubbles to analyze the role of such policy. He also questions the mechanism that an increase in interest rates actually deflates asset prices. The work of Assenmacher and Gerlach (2008) however shows using a VAR model that asset prices react almost instantaneously to the interest rate, which supports the conventional theory that interest rates matters through their role as the discount factor. Bernanke and Gertler (2000) use a model similar to mine where stock prices are represented as the price for capital, and bubbles are exogenous. They provide the theoretical benchmark result that asset price targeting is rather harmful in terms of welfare.

	$\pi$	$y$	$s$
$SD$	0.0092	0.0104	0.1407
$\pi$	1	-0.1734	-0.3867
$y$	-	1	0.6025
$s$	-	-	1

Table 2.1: Standard deviations and cross correlations of inflation, output and real stock prices, Core-Europe from 1976 to 2014 (quarterly). A detailed description of the data can be found in Section 2.3.

Episodes with booms and busts are recurrent phenomena. In analysing the housing market and equity prices in industrialized economies during the postwar period, the IMF (2003) found that booms in both markets arise frequently (on average every 13–20 years) with entailed drops in prices averaging around 30% and 45% respectively. These busts are associated with losses in output that

reflect declines in consumption and investment. Table 2.1 summarizes key statistics of European data and embedding a set of stylized facts:

- i) While standard deviations of inflation and output are roughly on the same level, the standard deviation of asset prices is roughly 10 times in magnitude.
- ii) Inflation is (weakly) countercyclical, i.e. negatively correlated with output.
- iii) Stock market prices and output are positively correlated.
- iv) Stock market prices and inflation are negatively correlated and as such more than inflation and output.

These stylized facts are well in line with empirical work, for instance by Campbell (1999) on stock prices and consumption. Barro (1990) and Sargent (2008) argue that stock prices are a good indicator for investment and hence future output, which implies a lead-lag structure of asset prices and output. Winkler (2014) conducts a vector auto-regression (VAR) on asset price shocks. He finds that the response of total factor productivity (TFP) is insignificant or even negative, while the asset price shock has significant effects on investment. Thus, he concludes that the classical view that stock price changes reflect new information about productivity changes might be controversial. Abbate et al. (2016) report similar findings by using a time-varying FAVAR. To my best knowledge, the relationship between inflation and stock prices has not been subject to detailed studies yet.

The model used here adopts the general New Keynesian literature (Woodford, 2003; Galí, 2008). Financial frictions are implemented inspired by BGG, but with the simplification that labor is the only input factor. This allows to focus on the dynamics of agents' interaction and interconnection of asset prices and macroeconomic activity. Boundedly rational expectations gained popularity for being able to endogenously reproduce asset price bubbles by specifying a behaviorally intuitive and simple mechanism of expectation formation. Empirical regularities of stock market prices such as return volatility (Shiller, 1981), return predictability (Fama and French,

---

<sup>1</sup>Monthly data from the OECD, stock prices are the MSCI-Europe index. Time series are HP-filtered with  $\lambda = 10e5$ . Prices in 2005, stock prices and output are divided by CPI.

1988) and fat tails of asset price distributions are hard to explain unless one relaxes the hypothesis of rational expectations. There is a growing literature on behaviourally rational agents that have been introduced to macroeconomic modelling, specifically with respect to expectations on output and inflation. For an extensive overview see e.g. Evans et al. (2001). Mankiw et al. (2003), Branch (2004), Pfajfar and Santoro (2008) and Pfajfar and Santoro (2010) provide empirical evidence in support of *heterogeneous* expectations using survey data on inflation expectations. Hommes et al. (2005), Hommes (2011), Pfajfar and Zakelj (2012) and Assenza et al. (2013) find evidence of simple, heterogeneous forecasting mechanisms in laboratory experiments with human subjects. Taken this evidence as a starting point, heterogeneous expectations have found their way into macroeconomic modelling, see for instance Evans and Honkapohja (2003), Tuinstra and Wagener (2007), Anufriev et al. (2008), Brazier et al. (2008), Branch and McGough (2009), Branch and McGough (2010), De Grauwe (2011), and De Grauwe and Macchiarelli (2013). From the perspective of this literature, recessions are not due to shocks to fundamentals but rather to massive coordination failure. Additionally, in other work (Boehl, 2017) I find that in markets with strong positive feedback, rational agents adopt the beliefs of boundedly rational traders and hence *act if* they were boundedly rational as well.

As to theoretical insights on the link between asset markets and real macroeconomic activity, Martin and Ventura (2011) rely on Gertler et al. (2010) to create a linkage between credit volume, firms' value and real activity and implement the idea of *rational bubbles* (Martin and Ventura, 2010). Tallarini (2000), Rudebusch and Swanson (2012) and others explore the dynamics implied by Epstein-Zin preferences (Epstein and Zin, 1989), while a different branch of the literature followed the idea of habit-formation specifications (e.g. Abel (1990), Ljungqvist and Uhlig (2000, 2015)). Also see Kliem and Uhlig (2016) for a brief overview of the literature and the estimation of such type of model. Generally, these methods report mixed success with fitting both, macroeconomic dynamics and asset price volatility. Greenwood and Shleifer (2014a) provide a summary on survey data that documents the failure of rational expectations. According to this data, the rational expectations hypothesis can, at least concerning asset markets, almost always be rejected.

The rest of this work is structured as follows. In Section 2.2 I formally present the macroeconomic part of the model and provide microfoundations for the mutual linkage between asset prices and macroeconomic aggregates. This model is used to study the equilibrium dynamics under rational expectations and to estimate and identify relevant parameter ranges in Section 2.3. I then endogenize fluctuations in asset prices in Section 2.4, where I also present simulation results and policy analysis. Section 2.5 concludes.

## 2.2 Model

The economy is populated by a continuum of identical households, a heterogeneity of firms, a financial intermediary and a monetary authority.

### 2.2.1 Households

Households are indexed by  $i$ . They face a standard problem of maximizing the expected present value of utility by deciding over consumption of a composite good  $C_t$  and time devoted to the labour market  $H_t$ . For each unit of labour supplied they receive the real wage  $W_t$ . Furthermore they can deposit monetary savings  $D_t$  at the financial intermediary. The maximization problem for individual agents is then

$$\max_{\{C_{i,t}\}, \{H_{i,t}\}, \{D_{i,t}\}} E_t \sum_{s=t}^{\infty} \beta^{s-t} \left( \frac{\zeta_{i,t} C_{i,s}^{1-\sigma}}{1-\sigma} - \xi \frac{H_{i,s}^{1+\gamma}}{1+\gamma} \right)$$

s.t. the budget constraint in real terms states

$$C_{i,t} + D_{i,t} \leq W_t H_{i,t} + R_t \frac{P_{t-1}}{P_t} D_{i,t-1} + \mu \int_0^{\bar{\omega}_t} \omega H_t / X_t dF(\omega) \quad \forall t = 1, 2, \dots$$

where  $\mu \int_0^{\bar{\omega}_t} \omega H_t / X_t dF(\omega)$  are the audition costs for defaulting wholesalers, which will be explained further below. Audition costs are, via the financial intermediary, distributed equally among households and do not enter optimality conditions. Each household is subject to an idiosyncratic



preference shock  $\zeta_{i,t}$ . The composite consumption good consists of differentiated products from the retail sector and is sold in a monopolistically competitive market. The composite good and the aggregate price index for the consumption good are defined by the CES aggregators

$$C_t = \left( \int_0^1 C_{l,t}^{\frac{\epsilon-1}{\epsilon}} \right)^{\frac{\epsilon}{\epsilon-1}} \quad \text{and} \quad P_t = \left( \int_0^1 P_{l,t}^{1-\epsilon} \right)^{\frac{1}{1-\epsilon}}$$

Optimization yields the usual Euler equation and a clearing condition for the labor market

$$\zeta_t C_t^{-\sigma} = E_t \left\{ \beta R_{t+1} \frac{P_t}{P_{t+1}} C_{t+1}^{-\sigma} \right\} \quad (2.2.1)$$

$$\xi H_t^\gamma = \frac{\zeta_t W_t}{C_t^\sigma} \quad (2.2.2)$$

where  $\zeta_t$  denotes the i.i.d. aggregate demand shock that is due to the idiosyncratic preference shocks. Since individual shocks are not observable for other agents, at time  $t$  the aggregate shock is not observable either. Given optimality, the budget constraint needs to hold as an equality and agents obey the transversality condition

$$\lim_{s \rightarrow \infty} \beta^{s-t} E_t C_s^{-\sigma} D_s = 0.$$

Household deposits are given to the sector of financial intermediaries.

## 2.2.2 Firms

To maintain analytical tractability, firms are divided into a wholesale and retail sector. Wholesalers borrow money from the financial intermediary to finance production and their shares are traded at the stock exchange. Their (homogeneous) good is sold to the retail sector where diversification takes place and the then heterogeneous goods are sold to the households with monopolistic profits.

## Wholesale Sector

Let labor be the only production factor and index wholesalers by  $j$ , then the CRS production function is

$$Y_{j,t} = \omega_{j,t} H_{j,t},$$

where  $\omega_{j,t}$  is a firm-specific idiosyncratic productivity shock similar to the households' preference shock. A more careful definition can be found in the appendix. The negative correlation between stock prices and inflation indicates that if stock prices are relevant for the macro economy, it should be motivated through the firms' financial structure. Therefore I specify two additional assumptions:

- i) Unlike creditors, shareholders can liquidate the firm at any time without costs.
- ii) Since that, wholesalers maximize the expected future stream of dividends on equity.

To simplify the optimization problem I allow for negative dividends to be paid, which implies that firms can obtain financing resources from their shareholders as well. Shareholders will comply if expected future profits are appropriately high. This is a reasonable assumption since shareholders are willing to increase firms' equity to seize the opportunity of higher future profits. A similar approach is chosen by Martin and Ventura (2010) for aggregated investment. Wholesalers are price takers. Let  $X_t$  be the gross markup of retail goods over wholesale goods. Then equivalently  $X_t^{-1}$  is the relative price of wholesale goods. This implies that  $R_t^H$ , the gross return on employing one unit of labor, is given by

$$R_{t+1}^H = (X_t W_t)^{-1}, \quad (2.2.3)$$

where the reciprocal definition of  $X_t$  ensures  $R_t^H > 1$ , a necessary condition for positive external finance. Let us denote firm  $j$ 's equity by  $N_{j,t}$ . The expected return on equity implied by the stock price,  $E_t R_{t+1}^S$ , is then  $\frac{R_{t+1} S_{j,t}}{N_{j,t}}$ , given no arbitrage. Goods are produced and sold in the current period, but returns are realized at the beginning of the next period. Then firms decide upon their equity and distribute the rest as dividends  $\Theta_{t+1}$ . Finally, the firm's shares are traded at the stock exchange. As shown further below, given no-arbitrage the price  $S_{j,t}$  of one share of the firm needs to satisfy

$S_t = E_t \frac{\Theta_{t+1} + S_{t+1}}{R_{t+1}}$  or equivalently

$$S_t = E_t \sum_{s=t}^{\infty} \prod_{l=t}^s R_{l+1}^{-1} \Theta_{s+1} \quad (2.2.4)$$

where the normalization of the number of shares to unity is implied. Dividends in period  $t$  are composed of  $\Theta_t = H_{t-1}/X_{t-1} - N_t$ , and  $N_t$  is the amount necessary to finance production costs  $N_t = W_t H_t$ .

**Simple example without external finance:** The link between stock prices and return per unit of labor  $R_{t+1}^H$  can be explained more intuitively in a world without external finance. Hence, let me briefly abstract from external finance. Recall that every period firms choose how much of their returns to retain and how much to distribute. The Lagrangian is

$$\max_{\{H_t\}, \{N_t\}, \{\lambda_t\}} E_t \sum_{s=t}^{\infty} \prod_{l=t}^s R_{l+1}^{-1} [H_{s-1}/X_{s-1} - N_s] - \lambda_s (W_t H_s - N_s).$$

The first-order condition is  $H_t/X_t = N_t R_{t+1}$ , which combined with the definition of expected future dividends gives  $E_t \Theta_{s+1} = R_{t+1} N_t - E_t N_{t+1}$ . Inserting this result into Equation (2.2.4) implies that stock prices reflects the value of equity perfectly as in

$$S_t = \frac{R_{t+1} N_t - E_t N_{t+1} + \frac{E_t \{R_{t+2} N_{t+1}\} + \dots}{R_{t+2}}}{R_{t+1}} = N_t.$$

It follows that the optimal labor demand  $H_t = R_{t+1} S_t X_t$  is determined by the prices prevailing in the financial market in combination with wholesale prices and the economies interest rate. Once we drop the assumption that expectations on financial market prices are perfectly rational, this can

lead to coordination failure.<sup>2</sup>

**Full model:** Let me now return to the wholesalers' problem with external finance. The volume of external finance demanded is firms' working capital  $W_t H_{j,t}$  minus equity, hence

$$B_{j,t} = W_t H_{j,t} - N_{j,t}. \quad (2.2.5)$$

Regarding the external finance premium I follow the lines of the BGG financial accelerator mechanism closely, where the borrowing process follows a costly state verification (CSV) approach. I follow a similar mechanism in the appendix to establish that the interest rate on loans from the intermediary, denoted by  $R_{t+1}^B$ , is a risk-premium on the prevailing interest rate which depends on the individual firm's leverage,

$$R_{j,t+1}^B = z \left( \frac{N_{j,t}}{W_t H_{j,t}} \right) R_{t+1}$$

with  $\frac{\partial z}{\partial N_t} < 0$ . Intuitively, when the leverage ratio decreases, the premium on external finance falls because the amount of collateral increases and the loan becomes less risky. I show in the appendix that optimality requires the return on assets to be equal to the rate paid on external funds,  $R_{t+1}^S = R_{t+1}^B$ . Otherwise wholesalers would have an incentive to increase or decrease the borrowing volume. Similarly to the example above, an increase in  $S_t$  will also have an increasing effect on equity  $N_t$ . Hence, once a functional form of  $z(\cdot)$  is known, we can use

$$\frac{S_{j,t}}{N_{j,t}} = z \left( \frac{N_{j,t}}{W_t H_{j,t}} \right)$$

---

<sup>2</sup> Another perspective is that the expected return implied by asset prices is

$$R_{t+1}^S = \frac{S_t}{N_t} R_{t+1}$$

which in combination with Equation (2.2.3) gives us the same pricing equation for the simplified problem without external finance. Note that without external finance  $N_t = H_t/X_t$ , i.e. equity and working capital are the same. Under rational expectations this implies  $\frac{S_t}{N_t} = 1$ .

to eliminate  $N_t$ . Since  $R_{t+1}^S = R_{t+1}^B$ , these returns also need to equal  $R_{t+1}^H$ . Plugging the result into Equation (2.2.3), substituting for  $W_t$  and  $H_t$  and log-linearising the result gives us an aggregate representation of the price  $X_t$  for wholesale goods:

$$x_t = -\eta y_t - i_t + v s_t.$$

A competitive market for wholesale goods implies equal prices. Likewise, all firms have to offer the same  $R_{t+1}^S$ . This implies that the stock market evaluation of shares determines the amount of equity and profits. In general equilibrium, to comply to market forces and implicit expectations on future dividends, relative prices have to rise. This mechanism can be summarized by *the pressure to perform* combined with the fact that managers can not distinct whether aggregate stock prices are overvalued or not.

Note that this result contains two effects: an increase in stock prices puts pressure on firms to increase prices. In an economy with constant marginal costs, firms' only chance to increase profits per labor unit, is to raise the price level. But an increase in stock prices also decreases the leverage ratio, which lowers the cost for external finance and lets prices decrease. In this model, the second effect prevails as intuition suggests.

## Retailers

Retailers buy the homogeneous good  $Y_{j,t}$  from entrepreneurs and differentiate to sell it in a monopolistic competitive consumer market. This implements the Calvo pricing mechanism as it is standard in the literature. For details on the solution given the markup  $X_t$ , see Bernanke et al. (1999). Letting resellers be denoted by  $l$ , it can be shown that setting the optimal price  $P_t^*$ , given the corresponding demand  $Y_{l,t}^*$ , satisfies

$$\sum_{k=0}^{\infty} \theta^k E_t \left\{ \Lambda_{t,k} \left( \frac{P_t^*}{P_{t+k}} \right)^{-\epsilon} Y_{l,t+k}^* \left[ \frac{P_t^*}{P_{t+k}} - \left( \frac{\epsilon}{\epsilon - 1} \right) X_t^{-1} \right] \right\} = 0. \quad (2.2.6)$$

Taking into account that each period the fraction  $\theta$  of retailers is not allowed to change prices in period  $t$ , the aggregated price level follows

$$P_t = [\theta P_{t-1}^{1-\epsilon} + (1-\theta)(P_t^*)^{1-\epsilon}]^{\frac{1}{1-\epsilon}}$$

where  $P_t^*$  needs to satisfy Equation (2.2.6). Log-linearizing the combination of both equations yields the Phillips Curve (2.3.1) depending on the log-linearised markup  $x_t$ .

### 2.2.3 Financial Intermediation

There is a continuum of financial intermediaries indexed by  $k$ . Each of them takes the deposits  $D_{k,t}$  received from households as given and invests a fraction in the stock market by holding a  $J_{k,t}$ -share of stocks at the real stock price  $S_t$  and issues the rest as credit volume  $B_{k,t}$  to the wholesalers. I assume that investment in the financial market is done by traders that are each associated with a financial intermediary. Furthermore the intermediary has access to central bank money for which he will have to pay the real central bank rate  $R_{t+1}$ . Next periods' real dividends net of seized collateral are expected to be  $E_{k,j}\Theta_{t+1}$ . Market clearing requires:

$$R_{t+1}^D D_{k,t} = \hat{E}_{k,t}[P_{t+1}\Theta_{t+1} + P_{t+1}S_{t+1}]J_{k,t} + z^{-1}R_{t+1}^B B_{k,t}$$

subject to the constraint  $D_{k,t} \geq P_t S_t J_{k,t} + B_{k,t}$ . From the fact that the opportunity costs of finance are given by the central bank interest rate, optimality requires  $R_{t+1}^D = z^{-1}R_{t+1}^B = \frac{\hat{E}_t\{P_{t+1}\Theta_{t+1} + P_{t+1}S_{t+1}\}}{P_t S_t} = R_{t+1}$ . In case of homogeneous and rational expectations the pricing equation for stocks can be aggregated straightforwardly:

$$R_{t+1}P_t S_t = E_t\{P_{t+1}\Theta_{t+1} + P_{t+1}S_{t+1}\}. \quad (2.2.7)$$

Let a capital letter without time subscript denote the respective steady-state value. In equilibrium,  $\Theta$  is a function of the markup  $X$  and total output  $Y$ .  $\frac{\partial \Theta}{\partial Y}$  depends on the share of labor income

and is here set to unity. Note that, when log-linearising equation (2.2.7)) the coefficient of  $(1 - \beta)$  of  $E_t y_{t+1}$  is very small, so introducing a more realistic labor income share would not be a notable improvement. The log-linear version of the asset pricing equation is thus almost independent of expectations on next periods' output and markup. This yields the following expression for the percentage deviation of stock prices from their steady state value, where  $r_{t+1}$  denotes the real interest rate  $r_{t+1} \equiv i_{t+1} - E_t \pi_{t+1}$ :

$$s_t = (1 - \beta)E_t y_{t+1} + \beta E_t s_{t+1} - r_{t+1}.$$

## 2.2.4 Central Bank and Government

The central bank follows a standard contemporaneous Taylor Rule. Additionally I allow to adjust the nominal interest rate  $i_{t+1}$  in response to the real stock market prices  $s_t$ :

$$i_{t+1} = \phi_\pi \pi_t + \phi_s s_t.$$

A policy that increases the nominal interest rate when stock market prices increase I will call *asset price targeting* (APT). Asset price targeting is my only additional policy measure that is explicitly implemented in the model. One problem the monetary authority faces when responding to movements in stock prices is that it is not ex-ante identifiable whether a deviation in asset prices represents a shift in fundamentals or in beliefs. In order to establish a practicable mechanism, if such policy is in place the central bank hence must *always* react to movements in stock prices (here in terms of deviation from steady state), independently of whether these are identified as bubbles or a correct anticipation of real future movements. I furthermore for reasons of tractability abstract from governmental expenditures here and assume that the government issues no debt.

## 2.3 General Equilibrium and Estimation

The linearized economy is characterised by the following set of equations:

$$\pi_t = \beta E_t \pi_{t+1} + \kappa x_t + v_t^\pi, \quad (2.3.1)$$

$$y_t = E_t y_{t+1} - \sigma^{-1} r_{t+1} + v_t^y, \quad (2.3.2)$$

$$x_t = \eta y_t + i_{t+1} - \nu s_t, \quad (2.3.3)$$

$$s_t = (1 - \beta) E_t y_{t+1} + \beta E_t s_{t+1} - r_{t+1} \quad (2.3.4)$$

$$i_{t+1} = \phi_\pi \pi_t + \phi_s s_t. \quad (2.3.5)$$

Shock terms and the real interest rate are given by

$$v_t^\pi = \rho_\pi v_{t-1}^\pi + \varepsilon_t^\pi, \quad \varepsilon_t^\pi \sim N(0, 0.1)$$

$$v_t^y = \rho_y v_{t-1}^y + \varepsilon_t^y, \quad \varepsilon_t^y \sim N(0, 0.1)$$

$$r_{t+1} = i_{t+1} - E_t \pi_{t+1}.$$

In this section it is assumed that all expectations are formed homogeneously and rationally. Equation (2.3.1) is the New-Keynesian Phillips curve which links inflation  $\pi_t$  to the markup  $x_t$ . Equation (2.3.2) is referred to as the dynamic IS-curve. Equation (2.3.4) states the no-arbitrage condition for the stock market. If stock prices do not feed back to the real economy, this equation would not contain any relevant information for macro dynamics. The major difference to the standard New Keynesian model thus lies in Equation (2.3.3) which implements the financial accelerator and outside-option effect discussed in Section 2.1. It imposes a relationship between output, markup, nominal interest rate and stock market prices. Stock market prices can be thought to act on the market return on investment and impact the firms' pricing decision, optimal leverage and, in turn, the external finance premium. This also establishes the linkage between the textbook model and the BGG-type credit frictions where  $\nu$  is the price elasticity of the markup with respect to stock market



prices.<sup>3</sup> Note that this does not nest the Woodford-Type model as a special case where  $\nu = 0$  since marginal costs would still depend on the nominal interest rate. For stock prices to be relevant for the macroeconomy it is hence necessary that  $\nu \neq 0$ , since otherwise this equation collapses to the markup  $x_t$  simply being a fraction of wages  $w_t$  and the interest rate. Equation (2.3.5) is the Taylor rule with inflation and asset price targeting.

$v^y \sim N(0, \sigma_y)$  represents the aggregate of individual preference shocks  $\zeta_t$  and translates to a demand shock. Since individual preferences are not publicly observable, the realization of the shock is not ex-post observable.  $v^\pi \sim N(0, \sigma_\pi)$  is an aggregate productivity shock that results from idiosyncratic productivity shocks to wholesalers. Similar as to the demand shock,  $v^\pi$  is not observable in the aggregate since it affects producers individually. Once I deviate from the assumption of rationality, the non-observability of both shocks is an important ingredient of my model. Both shocks follow an AR(1) structure with  $\rho_\pi$  and  $\rho_y$  respectively.

Equations (2.3.1) to (2.3.4) can be represented as a 3-dimensional system of the endogenous variables  $\pi_t$ ,  $y_t$  and  $s_t$ :

$$\underbrace{\begin{bmatrix} 1 - \phi_\pi \kappa & -\kappa \eta & \kappa \nu \\ \phi_\pi \sigma^{-1} & 1 & \phi_s \sigma^{-1} \\ \phi_\pi & 0 & 1 + \phi_s \end{bmatrix}}_{\mathbf{M}} \underbrace{\begin{bmatrix} \pi_t \\ y_t \\ s_t \end{bmatrix}}_{\mathbf{x}_t} = \underbrace{\begin{bmatrix} \beta & 0 & 0 \\ \sigma^{-1} & 1 & 0 \\ 1 & 1 - \beta & \beta \end{bmatrix}}_{\mathbf{P}} \underbrace{\begin{bmatrix} E_t \pi_{t+1} \\ E_t y_{t+1} \\ E_t s_{t+1} \end{bmatrix}}_{E_t \mathbf{x}_{t+1}} + \underbrace{\begin{bmatrix} v_t^\pi \\ v_t^y \\ 0 \end{bmatrix}}_{\mathbf{v}_t}. \quad (2.3.6)$$

The matrix  $\mathbf{N} = \mathbf{M}^{-1} \mathbf{P}$  and particularly the eigenvalues of  $\mathbf{N}$  take a non-trivial form if expressed as a mapping from the parameter space. Figures showing these eigenvalues are deferred to Appendix A.4.

---

<sup>3</sup> Note that (2.3.3) can also be interpreted as to incorporate the idea that creditors use stock prices as a proxy for firms' collateral. Given  $\bar{\nu}$  both mechanisms lead to the same model with only marginally different parameterization.

## Estimation and Identification

Most of the deep parameters are fixed to values that are commonly found in the literature. I let  $\beta = 0.99$ , representing the short-term perspective of a quarterly model and, in accordance, set the shocks' autocorrelation to  $\rho_\pi = 0.9$  and  $\rho_y = 0.7$  respectively. Other values are consistent with the calibration of Woodford (2003) as I pick  $\gamma = 0.3$  and  $\omega = 0.66$  to obtain  $\eta = \frac{\sigma + \gamma + \bar{\nu}}{1 - \bar{\nu}} \approx 1.58$  and  $\kappa = (1 - \omega)(1 - \beta\omega)/\omega \approx 0.179$ .  $\bar{\nu}$ , the elasticity of the external finance premium with respect to net worth determines the elasticity of marginal costs to changes in stock prices, is defined by  $\nu = \frac{\bar{\nu}}{1 - \bar{\nu}}$ . The central banks policy in the baseline setup is described by  $\phi_\pi = 1.3$  and  $\phi_s$ , the response in interest rate with respect to stock prices, is set to zero implying that the central bank does not target stock prices when setting the policy rate.

To identify parameter values for the benchmark model with rational expectations I make use of a grid-based minimization technique that is explained in detail in Appendix A.5. The underlying intuition of this technique is to find the parameters of the global minimum of a distance measure between the simulated moments and those presented in Table 2.1. For this exercise (as well as for the Bayesian estimation) for output and inflation I use European data in levels provided by the OECD.<sup>4</sup> while for stock prices I use the MSCI-Europe index.<sup>5</sup> These are then deflated by deviding by the price index. Each time series is in logs and HP-filtered with using  $\lambda = 1600$ . This data is quarterly and ranges from 03/1976 to 03/2015, hence a total of 158 observations is being used.

I furthermore also conduct Bayesian estimations, considering a cost-push shock as well as an exogenous shock on asset prices. These results are redirected to Appendix A.2. The first reason is, once the model is extend by a nonlinear speculative process the use of Bayesian estimation is not feasible since finding the likelihood function of a model potentially embedding deterministic components is highly nontrivial. Hence, if I want to create a benchmark calibration it is advisable to use the same calibration technique when identifying the rational expectations model. Secondly, Bayesian estimation targets more statistical moments than the ones considered here, namely the

---

<sup>4</sup> At the time of writing, this is available at <https://data.oecd.org/>.

<sup>5</sup> Downloaded from <https://www.msci.com/indexes>. Unfortunately the website does not support data exports anymore.

autocorrelation parameters of endogenous variables. Since the model presented here tends to be too simplistic to explain the data in richer detail, I rather target specific important moments than the whole spectrum and profit from the simplicity when identifying economic mechanisms. Furthermore, for reasonable values of risk aversion  $\sigma$  stock prices and output display a relatively similar response to changes in interest rates. Since the data correlation between stock prices and output is relatively strong, it is difficult to identify the direction of causality, i.e. whether a change in output is induced by asset prices that shift down marginal costs or directly by the interest rate setting. This leads to an overall poor identification of parameter values when using Bayesian estimation, which, regardlessly, confirms the parameterization above.

For the RE model I consider an additional cost-push shock on marginal costs  $x_t$  to provide an equal number of degrees of freedom as the extended model in the next section. Thus,

$$v_t^x = \rho_x v_{t-1}^x + \varepsilon_t^x, \quad \varepsilon_t^x \sim N(0, \sigma_x)$$

is added to Equation (2.3.3), which then reads

$$x_t = \eta y_t + i_{t+1} - \nu s_t + v_x.$$

Table 2.2 shows the parameter values estimated in the grid-based minimization procedure. Likewise

$\nu$	$\sigma_\pi$	$\sigma_y$	$\sigma_x$	$\rho_x$
0.22	0.001	0.026	0.002	.93

Table 2.2: Parameter estimates of the RE model using a grid-based minimizing procedure.

an exogenous shock on stock prices  $v_s$  also follows an AR(1) structure and hits the economy in the zero profit condition.<sup>6</sup> The central result is that  $\nu$  is significantly positive, implying that stock prices have a considerably strong impact on the price level via marginal costs. Note that deviations in stock prices in the data are approximately ten times as strong than deviations in inflation. A  $\nu$

---

<sup>6</sup> Standard deviations necessary to match the data are lower than in the Bayesian estimation since less moments are targeted.

of  $\approx .1$  then implies that the respective impact is large, and probably overestimated by the model. This is however useful for the scope of this paper since it provides a strong motivation for monetary policy. I will elaborate on this further below.

Table 2.3 reveals two key problems of the rational expectations based approach. First, as is well known, it is impossible to match the ratio of standard deviations properly. Asset prices are mainly driven by fluctuations in the interest rate, which in turn depend on deviations in inflation from the central banks target. Hence, unrealistically strong fluctuations in inflation are necessary to replicate the standard deviation of asset prices. Secondly, this leads to a very high correlation between inflation and stock prices, which is also not supported by the data. The only reason why stock prices and output are correlated is the existence of a feedback between stock prices and output, i.e. a positive  $\nu$ .

	$\pi$	$y$	$s$
$SD$	0.0642	0.0102	0.1257
$\pi$	1	-0.472	-0.999
$y$	-	1	0.501
$s$	-	-	1

Table 2.3: Standard deviations & cross correlations of the estimated Rational Expectation model. The model is unable to correctly capture the covariance between stock prices and inflation.

In the next section I will show how to significantly improve the data fit by introducing speculative dynamics.

## 2.4 Endogenous Fluctuations in Asset Prices

This section relates excess volatility in stock prices – a necessary feature to match the stylized facts from Table 2.1 – to expectations about stock prices. As argued in Section 2.1, bounded rationality is a natural candidate to explain the amplification of exogenous shocks. In line with the research on bounded rationality discussed earlier, I here drop the assumption that stock market expectations are fully rational and introduce speculative behavior at the financial market.

The fact that all markets other than the stock market are dominated by rational agent also preserves the forward looking nature of the model that comes along with the rational expectations structure. Let us denote the model consistent rational expectations on inflation and output by  $E_t \pi_{t+1}$  and  $E_t y_{t+1}$  using the rational expectations operator. Let speculative expectations on stock market prices be denoted by  $\hat{E}_t s_{t+1}$ .<sup>7</sup> This demands for a mechanism of how rational agents deal with the existence of agents that form different, non-rational beliefs. It is here assumed that the distribution of agent types is unobservable. Then, rational agents are ex-post unaware of the presence of non-rational agents.<sup>8</sup> Since aggregate shocks are unobservable as well, rational agents perceive fluctuations induced by speculation as part of this exogenous noise. Let  $\tilde{\mathbf{v}}_t$  be the *perceived* exogenous shocks, which, as I will show, depend jointly on the real exogenous shocks and the degree of financial market speculation.

Letting  $E_t [\mathbf{x}_{t+1} | \tilde{\mathbf{v}}_t]$  denote the rational expectations solution of (2.3.6) in terms of these perceived shocks, then I have to find a dynamic representation of

$$\mathbf{M}\mathbf{x}_t = \mathbf{P} \begin{bmatrix} E_t [\pi_{t+1} | \tilde{\mathbf{v}}_t] \\ E_t [y_{t+1} | \tilde{\mathbf{v}}_t] \\ \hat{E}_t s_{t+1} \end{bmatrix} + \mathbf{v}_t, \quad (2.4.1)$$

where  $\mathbf{v}_t$  denotes the actual stochastic shocks  $v_t^\pi$  and  $v_t^y$ . Let me assume rational agents are New-Keynesians and do not think that asset prices play a role, which is not only consistent with the vast majority of the literature but also an hypothesis that cannot be rejected by Bayesian estimation.

---

<sup>7</sup>  $\hat{E}_t$  here is rather a function than a mathematical operator.

<sup>8</sup> Boehl (2017) show that in a system where fully rational and boundedly rational agents coexist, the type of dynamics is similar but *even more unstable* than in systems where rational agents are unaware of the presence of boundedly rational agents.

The perceived law of motion for rational agents is

$$\begin{bmatrix} \mathbf{P} & \mathbf{0}_{3 \times 2} \\ \mathbf{0}_{2 \times 3} & \mathbf{I}_{2 \times 2} \end{bmatrix} E_t \begin{bmatrix} \mathbf{x}_{t+1} \\ \tilde{\mathbf{v}}_{t+1} \end{bmatrix} = \begin{bmatrix} \mathbf{M} & \mathbf{0}_{3 \times 2} \\ \mathbf{0}_{2 \times 3} & \boldsymbol{\rho} \end{bmatrix} \begin{bmatrix} \mathbf{x}_t \\ \tilde{\mathbf{v}}_t \end{bmatrix}, \quad (2.4.2)$$

where  $\boldsymbol{\rho}$  is a diagonal matrix containing the autocorrelation parameters and  $\mathbf{x}_t$  the vector of endogenous variables at  $t$ . This is a different way to write the system in (2.3.6) but with perceived exogenous shocks instead of the real exogenous shocks. In the appendix I derive this system and use eigenvector-eigenvalue decomposition to find the (linear) rational expectation solution to (2.4.2). Let this linear solution be denoted by the matrix  $\boldsymbol{\Omega}$ . It needs to hold by definition that

$$\begin{bmatrix} \pi_t \\ y_t \end{bmatrix} = \boldsymbol{\Omega} \begin{bmatrix} \tilde{v}_t^\pi \\ \tilde{v}_t^y \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} E_t[\pi_{t+1} | \tilde{\mathbf{v}}_t] \\ E_t[y_{t+1} | \tilde{\mathbf{v}}_t] \end{bmatrix} = \boldsymbol{\Omega} \boldsymbol{\rho} \begin{bmatrix} \tilde{v}_t^\pi \\ \tilde{v}_t^y \end{bmatrix}. \quad (2.4.3)$$

It follows directly that we can express the conditional expectations on inflation and output without explicitly solving for the perceived shocks  $\tilde{\mathbf{v}}_t$ :

$$\begin{bmatrix} E_t[\pi_{t+1} | \tilde{\mathbf{v}}_t] \\ E_t[y_{t+1} | \tilde{\mathbf{v}}_t] \end{bmatrix} = \boldsymbol{\Omega} \boldsymbol{\rho} \tilde{\mathbf{v}}_t = \boldsymbol{\Omega} \boldsymbol{\rho} \boldsymbol{\Omega}^{-1} \begin{bmatrix} \pi_t \\ y_t \end{bmatrix}$$

Let me plug-back this result for the expectations on output and inflation into Equation (2.4.1) and rearrange in terms of the actual exogenous states and speculative expectations on the stock market to obtain the actual law of motion. The latter can then be expressed as a mapping  $\boldsymbol{\Psi} : (\boldsymbol{\rho}, \boldsymbol{\Phi}, \boldsymbol{\phi}) \rightarrow \mathbb{R}_{3 \times 3}$ ,

where  $\Phi$  is the set of model parameters  $(\beta, \sigma, \nu, \eta, \kappa)$  and  $\phi$  the two policy parameters. Hence,

$$\begin{bmatrix} \pi_t \\ y_t \\ s_t \end{bmatrix} = \Psi \begin{bmatrix} v_t^\pi \\ v_t^y \\ \hat{E}_t s_{t+1} \end{bmatrix}.$$

This represents a solution for the rational expectations equilibrium in terms of the real shock terms with one degree of freedom, which is used for boundedly rational beliefs  $\hat{E}_t s_{t+1}$ . Note that this actual law of motion, by definition, is not known to any of the agents.

Recall that the Blanchard and Kahn (1980) condition for *determinacy* of the rational expectation solution requires that the number of forward looking variables coincides with the number of eigenvalues of  $\mathbf{N} = \mathbf{P}^{-1}\mathbf{M}$  that are lying outside the unit circle. Since the eigenvalues of  $\rho$  will always lie inside the unit circle, all three eigenvalues of  $\mathbf{N}$  need to be lying *outside* the unit circle. Using standard numerical calculus to determine the modulus of the eigenvalues of  $\mathbf{N}$  and the parameter values from Table 2.4, these numerical results can be found in the appendix. The system is *determinate* for values of  $\phi_\pi$  larger than one (which corresponds to standard findings in the literature). To further ensure determinacy,  $\phi_s$  is bounded by  $[-0.116, 0.465]$  with the interval decreasing in  $\nu$ .<sup>9</sup>

Which policy implications can we deduct from this model without further specification of an expectation formation mechanism? In the absence of real shocks the law-of-motion in (2.4.3) can be reduced to

$$\mathbf{x}_t = \Psi_{:,3} \hat{E}_t s_{t+1} \quad \text{and in particular} \quad s_t = \Psi_{3,3} \hat{E}_t s_{t+1}.^{10} \quad (2.4.4)$$

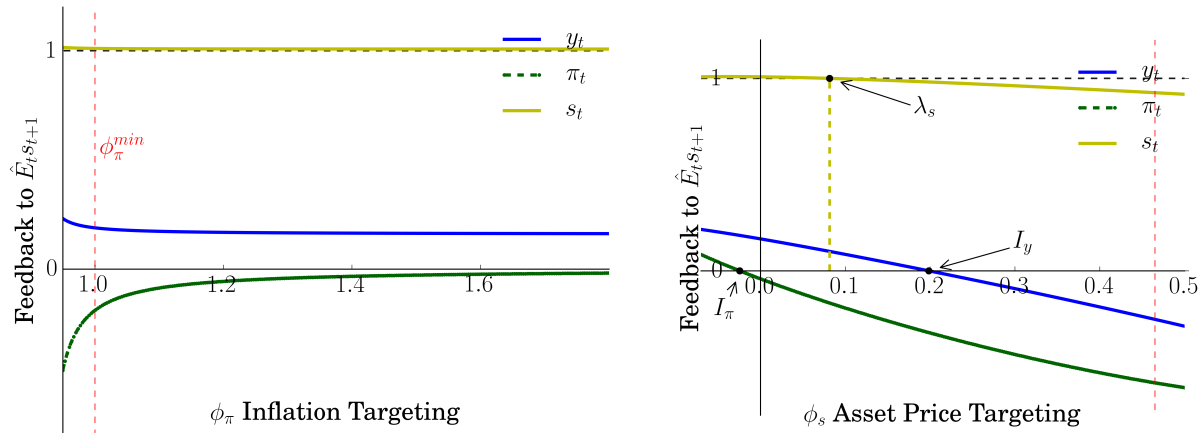
---

<sup>9</sup>Note that I explicitly do not rule out negative values of  $\phi_s$  to remain agnostic concerning an optimal policy. Naturally, these intervals are the boundaries for the following policy experiments, i.e. I am not considering rational sunspots in the expectations on inflation or output.

<sup>10</sup>I implement the model in Python. I would like to empathize the excellence of contemporary free and open source software, also and especially in comparison with proprietary software. I also want to encourage the reproducibility of research by providing the source code of my work which is available upon request. In the rest of this paper I use *Python-like* notation when referring to certain parts of matrices. So if

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix},$$

Let us consider the calibration from Table 2.2 and for now disregard any exogenous shocks in order to focus on economic intuition behind the impact of a deviation in stock price expectations.<sup>11</sup> Given this calibration  $\Psi_{3,3} > 1$  implying explosive expectations feedback in the financial market. Learning-to-forecast experiments and theoretical evidence<sup>12</sup> have shown that systems with positive feedback, especially when close to unit roots, can exhibit large swings and bubbles. Hence, when stabilizing such system it should *ceteris paribus* be the policy makers' aim to minimize  $\Psi_{3,3}$ . Likewise, the second best solution would be to minimize  $\Psi_{1,3}$  and  $\Psi_{2,3}$  and thereby minimizing the impact of stock prices on real activity. Following this line of argument,  $\Psi_{3,3}$  represents a key measure for the probability of excess volatility on the stock market.



(a)  $\Psi_{3,3}$  as a function of  $\phi_\pi$ , no asset price targeting ( $\phi_s = 0$ ).

(b)  $\Psi_{3,3}$  as a function of  $\phi_s$ ,  $\phi_\pi = 1.3$ .

Figure 2.1: General equilibrium effect of a 1% change in asset price expectations as a function of central bank policy parameters. Responses of output, inflation and asset prices in deviations from steady state.

Figure 2.1 shows the values in  $\Psi_{3,3}$  as a function of policy parameters. These can be interpreted as the general equilibrium response of endogenous variables to a one-percent increase in stock price expectations. An increase in  $\phi_\pi$  (left panel) has quite moderate impact on the system dynamics,

I use  $\mathbf{A}_{2:3,1:2}$  to denote the lower-left square matrix (row 2 to 3, column 1 to 2)  $\begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$  or  $\mathbf{A}_{3, \cdot}$  to denote the vector in the third row of  $\mathbf{A}$  given by  $[a_{31} \ a_{32} \ a_{33}]$ .

<sup>11</sup>  $\tilde{\gamma} = 0.23$  lies well in the 90% interval of the estimation and, compared to the value of 0.5 used in BGG, can be considered a conservative calibration.

<sup>12</sup> For a review on laboratory experiments on expectation formation, see Hommes (2011).



with almost flat curves in the relevant interval. Most noteworthy, the response of  $s_t$  with respect to  $\hat{E}_t s_{t+1}$  is almost constant in the degree of the policy response to the price level ( $\phi_\pi$ ), and always larger than one. However, the system changes more drastically with changes in  $\phi_s$ , with values of  $\Psi_{3,3} < 1$  for  $\phi_s > \lambda_s \approx 0.09$ . This means that if the central bank reacts moderately to stock prices, the positive feedback loop can be mitigated.

The diagram reveals two other interesting points. At  $I_y$  the impact of speculation on output is exactly offset and for higher values of  $\phi_s$  a positive shock on asset prices actually leads to a decrease in output. This is due to a stronger increase in the interest rate that dampens consumer demand. A similar point although for a negative value of  $\phi_s$ , is  $I_\pi$  where the inflation rate is completely unaffected by the immediate impact of a deviation in stock price expectations. We furthermore learn that fluctuations in asset price expectations can contribute to a high ratio of standard deviations of stock prices and output,  $\sigma_s/\sigma_y$ .

The above results imply that the central bank faces a trade-off. They can target to lower the impact of  $\hat{E}_t s_{t+1}$  on either  $\pi_t$  or  $y_t$ , but accept the extreme dynamic feedback induced by a high  $\Psi_{3,3}$  and its implications on the stability of the nonlinear expectations system. Or policy makers can choose to potentially stabilize the system but increase the impact of stock market expectations on real variables notably. After the analysis of the expectation-feedback system I will take a closer look at this trade-off. For this purpose it is necessary to analyse the explicit dynamics under speculation and to provide numerical results for the central bank policy, hence I implement a concrete mechanism on how financial market traders form expectations  $\hat{E}_t s_{t+1}$ .

### 2.4.1 Heterogeneous beliefs

Assume that traders follow the *Heterogeneous Agent Switching Model* (Brock and Hommes, 1998) and they switch endogenously switch between simple forecasting heuristics. This model provides a set of empirically relevant properties, most notably by incorporating the fact that agents can run short-term profits by following simple prediction strategies that were successful in the recent past, and these traders are able to outperform others that believe that the price will return to the

rational expectations equilibrium. It also embeds a nonlinear law of motion for asset prices that in particular can reproduce a positive correlation between returns and expected returns, and fat tails of the distribution of asset prices. For an intuition of the model dynamics and empirical validation see Hommes (2006).

Traders are heterogeneous in their forecasting rules. Let there be  $H > 1$  predictors of future prices and let each predictor  $h = 1, 2, \dots, H$  be of the form  $\hat{E}_{t,h} s_{t+1} = g_h s_{t-1} + b_h$ . I aggregate over each individual optimality condition (Equation 2.2.7) to derive the economy wide price for shares  $S_t$ .<sup>13</sup> If I let  $n_{t,h}$  denotes the fraction of traders using predictor  $h$  at time  $t$ , then

$$R_{t+1} \hat{E}_t \left\{ \frac{P_t}{P_{t+1}} \right\} S_t = E_t \Theta_{t+1} + \sum_h n_{t,h} \hat{E}_{t,h} S_{t+1}.$$

Let us also assume that traders take the real interest rate  $r_{t+1}$  as given. Log-linearization again yields

$$s_t = \beta \hat{E}_t s_{t+1} - r_{t+1} \quad \text{with} \quad \hat{E}_t s_{t+1} = \sum_h n_h \hat{E}_{h,t} s_{t+1}. \quad (2.4.5)$$

Not surprisingly, the first part here is identical to Equation (2.2.3) but the second part incorporates the speculative expectations  $\hat{E}_t s_{t+1}$ . Note that  $r_{t+1}$  in turn depends on  $y_t$  and  $\pi_t$ , so this equation yet takes the general equilibrium effect of changes in stock market prices into account. Fractions  $n_{h,t}$  are updated according to a *performance measure*  $U_{h,t}$  of predictor  $h$  in period  $t$ . As such I chose *realized past profits* as in

$$U_{h,t} = (\beta s_t - s_{t-1})(\beta \hat{E}_{t-1,h} s_t - s_{t-1}). \quad (2.4.6)$$

The choice of the performance measure is an essential ingredient of the model. It determines the properties of the nonlinear part of the dynamic system. Realized profits from trading qualify in several ways for our purpose. As outlined above, the fundamental difference between macroeco-

---

<sup>13</sup>I assume that trader  $k$ 's demand for shares is a linear function of expected profits with some  $\tau$ :

$$J_{k,t} = \tau \hat{E}_{k,t} \{ \Pi_{t+1} + S_{t+1} - R_t \frac{P_t}{P_{t+1}} S_t \}$$

with  $\int_0^1 j_t^k dk = 0$  when expressed as log-deviations from steady state.

nommic real markets and stock markets is that participants can make profits from speculation. Instead of being rewarded for an accurate estimate of the price, it is sufficient to forecast the direction of a price change correctly, hence to decide whether to go short or long. Likewise, a trader  $A$  that has a high forecast of next periods' prices will invest more money in the asset than some trader  $B$  with a relatively lower forecast. If it then turns out that  $B$  was correct in terms of point estimates,  $A$  will still realize higher profits since he invested more. This feature is captured by Equation (2.4.6).<sup>14</sup>

The probability that predictor  $h$  is chosen is given by the *multinomial discrete choice model*

$$n_{h,t} = \frac{e^{U_{h,t-1}}}{Z_{t-1}} \quad \text{and} \quad Z_{t-1} = \sum_{h=1}^H e^{U_{h,t-1}}. \quad (2.4.7)$$

As for the different predictors, consider a simple 3-type model where one type of agents are fundamentalists and the other two share a trend-following parameter  $\gamma$  and are either negatively or positively biased by  $\alpha$  as in

$$\begin{aligned} \hat{E}_{t,1}s_{t+1} &= 0, \\ \hat{E}_{t,2}s_{t+1} &= \gamma s_{t-1} + \alpha, \\ \hat{E}_{t,3}s_{t+1} &= \gamma s_{t-1} - \alpha. \end{aligned} \quad (2.4.8)$$

Now that expectation formation mechanisms for both types of agents are given, the model is fully specified. It consists of a linear part, associated with the economy and the formation of rational expectations and represented by Equation (2.4.3), and a nonlinear mechanism for boundedly rational expectation formation given by  $\hat{E}_t s_{t+1}$  and the performance measure  $U_{h,t}$  (Equations 2.4.5 and 2.4.6), the fractions  $n_{h,t}$  and the normalization factor (Equation 2.4.7), and the predictors (Equation 2.4.8).

---

<sup>14</sup>  $\tau$ , the demand sensitivity with respect to risk aversion, runs together with the sensitivity of choice, which normally is an integral part of the model. Since the intensity of choice does not have a measurable empirical counterpart, the product of both parameters is normalized to one here. This is to limit the degrees of freedom of the estimation, furthermore the other behavioral parameters can compensate this limitation fairly well. Including the nominal interest rate  $R_t \frac{P_t}{P_{t+1}}$  in the performance measure does not change the dynamics in a fundamental way, but leads to a slight asymmetry of bifurcations. Then it can not be guaranteed anymore that the mean of the time series of *perceived* shocks equals zero. This, however, is a necessary requirement when solving for rational expectations.

$\nu$	$\sigma_\pi$	$\sigma_y$	$\gamma$	$\alpha$
.10	.001	.007	1.31	1.18

Table 2.4: Parameter estimates for the model with endogenous dynamics in stock prices, based on a grid-based minimization procedure.

	$\pi$	$y$	$s$
$SD$	0.016	0.013	0.125
$\pi$	1	-0.176	-0.513
$y$	-	1	0.667
$s$	-	-	1

Table 2.5: Standard deviations and cross correlation matrix for simulations (10.000 iterations) of the model with endogenous dynamics in stock prices.

The remaining task is to assign values to the set of parameters for the boundedly rational expectation formation mechanism,  $\{\gamma, \alpha\}$ . Since it is not possible to use standard optimization techniques to find the parameter setup with the best fit, I use a grid method in combination with brute-force computation to find the set of behavioral parameters together with the standard deviations of shocks and the elasticity to stock prices. Details on the optimization algorithm can be found in Appendix A.5. The relevant parameters for cross correlation and standard deviations are thus only the parameters  $\nu$ ,  $\sigma_y$ ,  $\sigma_\pi$  and amplification process of stock prices, which emerges from the speculative process. The behavioral parameters can also be fine tuned to match turning points and amplitude. Optimal parameter values are summarized in Table 2.4. The simulated statistical moments can be found in Table 2.5.

I will provide economic intuition on why the extended model matches the data well after explaining the intuition behind this model in the following subsection.

## 2.4.2 Deterministic Simulations

To provide intuition behind the speculative process and the associated macroeconomic dynamics I use bifurcation theory. This also enables to study the dynamic properties of the system and to identify the relevant types of dynamics that occur given different values for the set of behavioral parameters  $\{\gamma, \alpha\}$ . For this purpose shocks are switched off and I run 11.000 iterations for each

combination in the parameter space of which I omit a transition phase of 10.000. All points of endogenous parameters visited in the remaining 1000 simulations are then plotted for each point in the parameter space. The relevant dynamics can be summarized by a bifurcation diagram. The diagram for  $\alpha$  is shown in Figure 2.2 and hence depicts the long-run dynamics as a function of the parameter.<sup>15</sup>

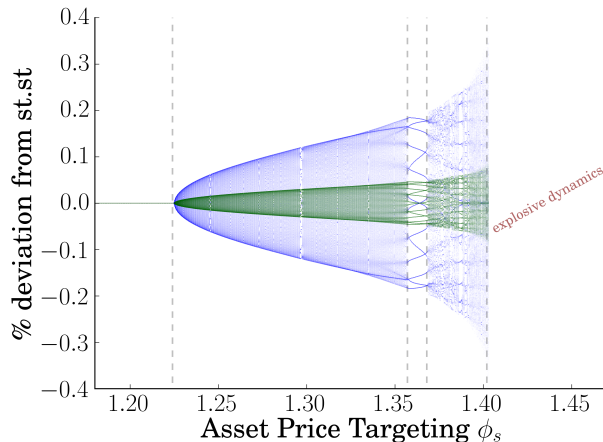


Figure 2.2: Long-run deterministic dynamics of  $y_t$  (blue/light) and  $\pi_t$  (green/dark) with respect to the behavioral bias  $\alpha$ .  $\gamma = 1.29$ , all other parameters as in Table 2.4.

For illustrative purposes value of  $\gamma$  is chosen for these figures which is slightly lower than the value identified by the estimation procedure.<sup>16</sup> Generally, an increase in the behavioral bias  $\alpha$  implies two effects. First the *quantitative* aspects of the dynamics change, i.e. the standard deviation increases. Secondly, the *type* of dynamics changes, each which is indicated by the vertical grey lines. From Figure 2.2 we learn that for low values of  $\alpha < 1.22$  the fundamental steady state is stable and unique, implying that the speculative forces are not strong enough to have an impact on stock prices without exogenous shocks. A stable steady state implies that there is always a fraction of fundamentalists in the market that outweighs the beliefs of biased agents. Exogenous shocks could lead to a temporal increase in the fraction of belief-biased agents, but their belief is yet not

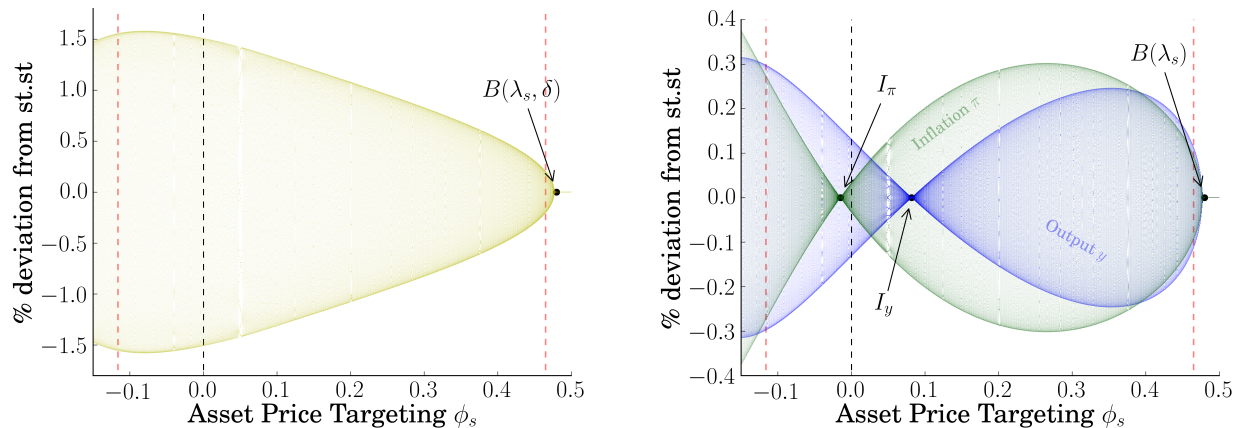
<sup>15</sup> As a technical note, it is clear that the first-order Taylor approximation around the steady state, that is embedded in the baseline DSGE model, will not hold once stock market prices deviate too much from their steady state value. I accept this inconvenience for the sake of simplicity and argue that introducing further nonlinearities would further complicate the interpretation of results while the added value of such undertaking would be unclear.

<sup>16</sup> For higher values of  $\gamma$  the type of dynamics are the same, but the interval of  $\alpha$  in which they occur is very small. The dynamics then change from stable steady state to explosive almost immediately when increasing  $\alpha$ .

strong enough to prevent the price from returning to its steady state. When  $\alpha$  increases limit cycles arise with the amplitude increasing with the parameter value. A small deviation from the steady state is now, by increasing the fraction of belief-biased agents, strong enough to ignite deterministic dynamics. Since beliefs are self-fulfilling to a large extent, the fraction of belief-biased agents is increasing each period, which in turn lets the price rise even further. When the price approaches the value predicted by positively (negatively) biased traders, they reduce their long (short) position. This then reduces their profit and other strategies become more attractive. Once alternative beliefs are more and more enforced, the fraction of positively (negatively) biased traders is too small to maintain the high price level. These financial cycles now have impact on output and inflation, which adapt the cyclic movement of stock prices as shown in the graphic. After  $\alpha \approx 1.35$  cycles become unstable and for values below  $\alpha \approx 1.37$  the simulations suggest that the system is close to a homoclinic orbit: the zero steady state is globally stable but locally unstable (Hommes, 2013). Long periods of stability can then be interrupted by busts in asset prices that are hard to predict, and, through the credit-collateral channel, can be followed by severe recessions. When  $\alpha$  increases even further cycles collapse and dynamics become explosive since the fraction of fundamentalists is not sufficiently large to stabilize the system ( $\alpha$  higher than 1.4). Since agents continue to assume price increases in the next period, this process is self-fulfilling and prices explode to infinity in the long run.

The data fitting algorithm suggest a value of  $\alpha$  of 1.18, which lies in the region where the deterministic steady state is stable. This result implies that the data is not driven by endogenous dynamics, but that the speculative dynamics are sensitive to the exogenous economic shocks and can induce excess volatility. Triggered by a series of exogenous shocks, agents might observe the change in stock prices and extrapolate it, follow their beliefs but then, caused by another exogenous shock, such bubbles might burst unexpectedly while preserving the mean-reverting property of a rational expectations model. This endogeneity also ensures that the correlation between stock prices and inflation is not overestimated as in the rational expectations model. Since the impact of an increase in stock prices decreases inflation through the marginal cost channel, the central bank

lowers the interest rate which in turn stimulates demand. This ensures that the correlation between output and stock prices is positive and noticeable, which would not be the case in a model without a feedback from stock prices to real activity. Hence, the property of excess volatility in combination with a mutual linkage is crucial to replicate key-moments of the data.



(a) Bifurcation diagram of asset prices  $s_t$  with respect to  $\phi_s$ .

(b) Bifurcation diagram of  $y_t$  (blue/light) and  $\pi_t$  (green/dark) with respect to  $\phi_s$ .

Figure 2.3: Deterministic dynamics for  $\phi_s$ .  $\alpha = 1.31$  and  $\gamma = 1$ , all other parameters as in Table 2.4.

### 2.4.3 Policy and Stochastic Simulations

To understand the interplay of monetary policy and speculative stock prices let me set  $\alpha$  to a higher value than implied by the estimation and likewise decrease  $\gamma$ , which helps us to understand what would happen in a world that would mainly be driven by speculative dynamics.<sup>17</sup> Still I abstract from stochastic shocks which means that yet, all fluctuations are endogenous.

As shown in the previous Subsection an increase in  $\phi_s$  is able to mitigate the impact of stock prices at least on output and decreases the positive feedback within the speculative process. Figure 2.3 show the long-run system's dynamics as a function of the policy parameter  $\phi_s$ . This diagram reveals that an *increase* in  $\phi_s$  has the same dynamic implications as a *decrease* in  $\alpha$ . The left figure

<sup>17</sup> Similarly to the long run dynamics depending on  $\alpha$ , the calibration implied by the data suggests that dynamics jump from stable steady state to explosive with only a marginal change in  $\phi_s$ .

suggests invariant cycles for all  $\phi_s < B(\lambda_s)$ . Since the dynamic process for  $s_t$  depends crucially on  $\Psi_{3,3}$ , the amplitude decreases with the magnitude of policy  $\phi_s$  and the graph exhibits a supercritical Hopf-Bifurcation in  $B(\lambda_s)$ , after which the dynamics settle to the steady state. The decrease in amplitude can be explained by the central banks response to stock prices. If  $\phi_s = 0$  an increase in asset prices decreases marginal costs and induces a drop in inflation, which is counteracted by an decrease of the nominal interest rate. Falling interest rates however increase stock prices even further. If the central bank now increases the interest rate when stock prices go up, this counteracts the first effect of a decrease in interest rates and dampens the positive feedback of stock price expectation to stock prices, which is captured by  $\Psi_{3,3}$ . The weaker this feedback, to a lesser are speculative dynamics self-fulfilling. In my example calibration depicted here, this leads to the fact that endogenous, deterministic cycles can be completely switched off at  $B(\lambda_s)$ . Under this parametrization a very high central bank policy with respect to stock prices would manage to shift the system on a stationary steady state path.

The dynamics for  $y_t$  and  $\pi_t$  are a linear mapping of the dynamics of  $s_t$  through  $\Psi$ . Both mappings for  $y$  and  $\pi$  contain an inflection point  $I_y$  and  $I_\pi$  respectively. These points correspond to those marked in Figure 2.1b. Since the results depicted here are deterministic and only result from speculative dynamics, this means that at these points the direct impact of speculation and bounded rationality can be completely offset. As explained before, the initial response of the central bank is to decrease the interest rate. If however  $\phi_s = I_y$ , this is completely mitigated by the *increase* of interest rates in response to stock prices, i.e. the net change in interest rates is zero. Since in this simple model output dynamics work through intertemporal substitution, output will not deviate from its steady state level. Likewise, if the central bank slightly decreases the interest rate when stock prices increase, this raises demand and, at point  $I_\pi$ , can perfectly set off the negative effect of stock prices on marginal costs. Hence at  $I_\pi$  the net change in marginal costs is zero and, through the Phillips curve, inflation remains at its target level.



The parameters in Table 2.4 provide a good data fit, have a meaningful behavioral interpretation and reproduce dynamics that match the stylized facts. Let me now use the parameters as the starting point to run stochastic simulations on the policy parameter  $\phi_s$ .

In Figure 2.4 standard deviations of simulations are shown where the real economic shocks  $v^\pi$  and  $v^y$  are added. Remember that under this parameterization, in the absence of shocks there would be no deterministic fluctuations. Accordingly the dynamics are a combination of stochastic shocks and endogenous responses of financial market speculation to these shocks. The implications from the deterministic case however still hold: an increase of  $\phi_s$  to a value higher than  $I_y$  induces unnecessary fluctuations in both inflation and output. In addition to that, the policy response to asset prices also reacts to movements in stock prices that are not induced by speculation but by the stochastic process. For a productivity shock this means that when inflation increases, the Taylor rule implies an increase in interest rate. This in turn will deflate stock prices. But if the central bank also targets stock prices, in equilibrium it will again lower the interest rate, shifting it away from the optimal response to the productivity shock which then overall induces an unnecessary strong response in inflation and output. The response to a demand shock works similarly, implying that asset price targeting always increases volatility in inflation and output. This effect runs in the opposite direction of the stabilizing effect of asset price targeting on speculative dynamics and the respective spillovers to inflation and output. This means that the optimal sensitivity of monetary policy to asset prices is bounded by  $I_\pi$  and  $I_y$ .

Furthermore, as predicted by the deterministic setup, a policy that *lowers* the interest rate in response to stock price increases also intensifies the feedback loop. Given the parameterization here, this leads to explosive dynamics in the stock market that are transmitted to the real economy. Monetary policy faces a trade-off that can be summarized by “*fragility versus volatility*”: first, raising  $\phi_s$  will decrease the volatility of  $s_t$  while keeping the volatility of output at a similar level, but increases the volatility of  $\pi_t$ . Although raising  $\phi_s$  mitigates the impact of speculation on real variables and strengthens the overall stability of the system, under the parametrization identified by the estimation procedure this stabilizing effect is of second order.

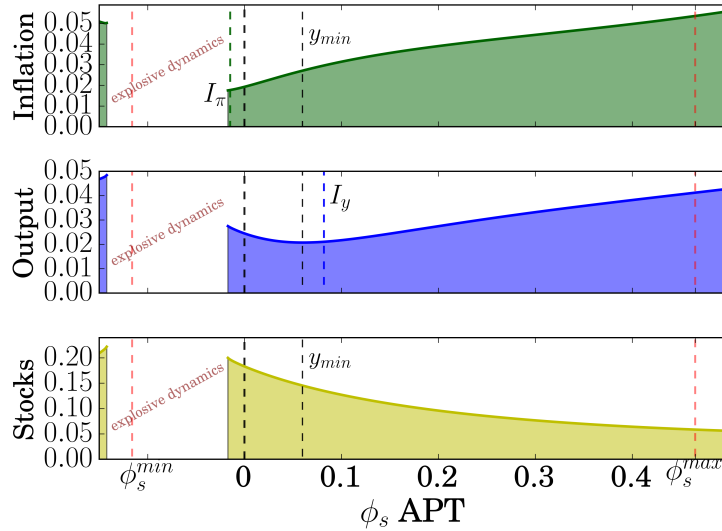


Figure 2.4: Standard deviations of stochastic simulations for given the asset price targeting policy parameter ( $\phi_s$ ) and the model with endogenous dynamics in stock prices. The economy is driven by exogenous shocks, but the nonlinear process in the stock market leads to endogenous amplification.

The model presented does not motivate the inclusion of asset prices themselves into welfare considerations. Relevant for the wellbeing of households is only their impact on output and inflation. Every welfare or central bank loss function that embeds a relative ranking of alternative central bank policies can be expressed as a convex combination of these two variances. From the point  $y_{min}$  on, both the variance of output and inflation are quasi-monotonous in  $\phi_s$ . It can therefore be concluded that only a very moderate asset price targeting has the potential to increase dynamic stability, decrease volatility and to mitigate the coordination failure induced by speculation in the financial market. Given the model presented here and the simulations depicted in Figure 2.4 it is unlikely that even such a moderate policy will have a positive effect, neither on welfare nor on financial stability. A concrete advice for policy however depends on how much a central bank weights fluctuations of output when conducting monetary policy.

This setup can also be used to analyze the policy implications given by the literature on rational bubbles, as proposed most prominently by Galí (2013). Since such rational bubble would presumably grow proportional to the interest rate, it is suggested to actually lower the policy rate when facing asset price bubbles. If such policy is non-discretionary, in my model it would lead

to a considerable increase in output and stock price volatility and, even for small values of the sensitivity of such policy, would lead to explosive dynamics in the stock market.

## 2.5 Conclusion

In this work I show that a causal linkage from stock prices to real activity as well as speculation in the stock market can help to replicate key-moments of the empirical data on inflation, output and stock prices. In particular the covariances between stock prices and inflation and stock prices and output, as well as the relatively high standard deviation of stock prices are well explained by the excess volatility that is induced by financial market speculation.

Methodologically, by arguing that financial markets entail a list of idiosyncracies, I introduce speculative behavior in the asset markets. Given a small number of consistency assumptions I show that it is possible to find the rational expectation solution on all other markets even though expectation formation in the asset market is of a boundedly rational type. Graphical evidence and simulation results suggest that any kind of speculation or herding, induced by bounded rationality and/or speculative profits, destabilizes the economy. Depending on the parametrisation, financial market interactions can lead to large and persistent booms and recessions. I thus find that instability is an inherent threat to economies with speculative financial markets.

I provide theoretical evidence that the central bank's interest rate setting amplifies the expectation feedback in the financial market, and that this can lead to an overall unstable dynamic process. My estimation procedure identifies the link from asset prices to real aggregates to be small but of macroeconomic significance, which implies a potential role for macroprudential policy to mitigate the negative externalities running through this mechanism.

Using this model as a counter-factual I show that if stock prices impact macroeconomic aggregates, a monetary policy rule that also targets asset prices can mitigate the excess volatility of stock prices but at the cost of intensifying real shocks. The model however suggests that such policy is bounded very narrowly by the unwanted side effects (“collateral damage”) of asset price

targeting. This result does not only hold for a policy that raises the interest rate when facing stock price bubbles but also for a policy that lowers the interest rate. The latter can easily amplify the speculative process and destabilize the economy furthermore.

Apart from the theoretical results provided in this paper, policy institutions may be well-advised to handle tools such as asset price targeting with care since such instruments might add a structural link between asset prices and macroeconomic aggregates. Such additional link embeds the risk of other unforeseeable complications, independently of how tight the natural link from asset prices to real activity is. This is particularly true because stock prices impact solely through signaling effects. Then an artificial inflation of stock prices in recessions might be counterproductive. This however leaves room for other macroprudential policies that potentially limits speculation itself or reduces profits from speculation in financial markets (i.e. policies such as short-selling constraints or leverage requirements) and indicates that such policy would contribute to overall economic stability. This work also suggests that neither stock prices nor indices on stock prices are a good indicator to base decisions on (lending, evaluation of competitors, . . .). Furthermore practitioners should be aware that, regarding stock prices and real activity, causality might run in both directions.

# Appendix A

## A.1 Entrepreneurs' optimization problem

I follow Bernanke et al. (1999) closely, but instead of assuming risk in the productivity of capital I assume idiosyncratic risk in labor productivity, i.e. firm  $j$ 's ex post gross return on one unit of labor is  $\omega_j$  which is i.i.d. across time and firms with a continuous and once-differentiable c.d.f  $F(\omega)$  over a non-negative support and with an expected value of 1. I assume that the hazard rate  $h(\omega) = \frac{dF(\omega)}{1-F(\omega)}$  is restricted to  $h(\omega) = \frac{\partial(\omega h(\omega))}{\partial \omega} > 0$ . The optimal loan contract between financial intermediary is then defined by a gross non-default loan rate,  $Z_{j,t+1}$ , and a threshold value  $\bar{\omega}_{j,t}$  of the idiosyncratic shock  $\omega_{j,t}$ . For values of the idiosyncratic shock greater or equal than  $\bar{\omega}_{j,t}$  the entrepreneur will be able to repay the loan, otherwise he will default.  $\bar{\omega}_{j,t}$  is then defined by

$$\bar{\omega}_{j,t} R_t^H H_{j,t} = Z_{j,t+1} B_{j,t}.$$

Dropping firms' subscripts, as in Bernanke et al. (1999) the optimal contract loan contract must then satisfy

$$\left\{ [1 - F(\bar{\omega}_t)] \bar{\omega}_t + (1 - \mu) \int_0^{\bar{\omega}_t} \omega dF(\omega) \right\} H_t / X_t = R_t (W_t H_t - N_t),$$

and the expected return to the wholesaler is (dropping time-subscript of  $\omega_t$  for better readability)

$$E \left\{ \int_{\bar{\omega}}^{\infty} \omega dF(\omega) - (1 - F(\bar{\omega})) \bar{\omega} \right\} H_t / X_t.$$

Given constant returns to scale, the cutoff  $\bar{\omega}$  determines the division of expected gross profits  $H_t/X_t$  between borrower and lender. Let me define

$$\Gamma(\bar{\omega}) = \int_0^{\bar{\omega}} \omega f(\omega) d\omega - \bar{\omega} \int_{\bar{\omega}}^{\infty} f(\omega) d\omega,$$

the expected gross share of profits going to the lender with  $\Gamma'(\bar{\omega}) = 1 - F(\bar{\omega})$  and  $\Gamma''(\bar{\omega}) = -f(\bar{\omega})$ .

This implies strict concavity in the cutoff value. I define similarly the expected monitoring costs as

$$\mu G(\bar{\omega}) = \mu \int_0^{\bar{\omega}} f(\omega) d\omega,$$

with  $\mu G'(\bar{\omega}) = \mu \omega f(\omega)$ . See BGG for the proof that the following result is a non-rationing outcome. The resellers problem of choosing the optimal equity can be solved by maximizing discounted profits over equity, or maximizing return on investment and including investment as part of the optimization problem.<sup>1</sup> Thus

$$\begin{aligned} \max_{\{H_t\}, \{\bar{\omega}_t\}, \{N_t\}, \{\lambda_t\}} E_t \sum_{s=t}^{\infty} N_t^{-1} \prod_{l=t}^s R_l^{-1} [(1 - \Gamma(\bar{\omega}_s)) H_s/X_s - N_{s+1}] \\ - \lambda_s ([\Gamma(\bar{\omega}_t) - \mu G(\bar{\omega})] H_t/X_t - R(W_t H_t - N_t)). \end{aligned} \quad (\text{A.1.1})$$

The first-order conditions for this problem can be written as:

$$\begin{aligned} H &: (1 - \Gamma(\bar{\omega}_t)) (X_t N_t R_t)^{-1} - \lambda_t ([\Gamma(\bar{\omega}_t) - \mu G(\bar{\omega})] / X_t - R_t W_t) &= 0 \\ \bar{\omega} &: \Gamma'(\bar{\omega}_s) (N_t R_t)^{-1} - \lambda_t [\Gamma'(\bar{\omega}_t) - \mu G'(\bar{\omega})] &= 0 \\ N &: -\frac{S_t}{R_t N_t^2} - R_t \lambda_t &= 0 \\ \lambda &: [\Gamma(\bar{\omega}_t) - \mu G(\bar{\omega})] H_t/X_t - R(W_t H_t - N_t) &= 0 \end{aligned}$$

---

<sup>1</sup> In fact the linkage between stock and wholesale prices would be clearer if I would allow wholesalers to have some monopolistic power and hence scope to adjust prices in response to pressure from the financial market. I do not model this here in detail since it seems unnecessary given the fact that the monopolistic competition is already implemented in the retail sector and the artificial division between both sectors is only due to reasons of analytical tractability.

Combining the first three conditions implies a connection between the optimal choice of labor, prices and stock prices. Using the optimality condition for the cutoff value  $\bar{\omega}_t$  and rearranging yields

$$\frac{\Gamma'(\bar{\omega}_t)}{\Gamma'(\bar{\omega}_t) - \mu G'(\bar{\omega}_t)} = \frac{S_t}{R_t N_t}$$

where the LHS can be written as a function  $\rho(\bar{\omega})$ . Bernanke et al. (1999) show that under reasonable assumptions  $\rho(\bar{\omega})$  is a mapping from  $\bar{\omega}$  to  $\mathbb{R}^+$ . The inverse of  $\rho(\cdot)$  can be used to establish that the premium payed on external funds depends on the return payed on internal funds. As noted in the main body, this is intuitive since the marginal costs of external and internal finance need to be equal. Likewise the risk premium on external funds can be defined to be a function of the leverage ratio (if  $N_t = W_t H_t$ , the premium is obviously one), which establishes the relationship in the main body.

## A.2 Bayesian estimation

Similar to the cost-push shock from the main body, here an exogenous shock on stock prices  $v_s$  is also considered. This process also follows an AR(1) structure and hits the economy in the zero profit condition. Note that such a third shock is also necessary to avoid stochastic indeterminacy. For the Bayesian estimation I make use of the standard routines implemented in Dynare (Adjemian et al., 2011).

Priors and the result of the estimation can be found in Table A.1. As in Section 2.3, most priors are taken from Smets and Wouters (2003) while the value of the prior for  $\gamma = 0.3$  is consistent with the calibration in BGG. In order to remain agnostic about the value of  $\nu$ , its prior is chosen quite broadly with a mean equal zero, while this value is set to 0.5 in BGG. I use a prior of  $\phi_\pi = 1.3$  as the central bank's policy parameters. Note that the prior for  $\sigma$  is slightly tighter than in Smets and Wouters (2003). This is necessary for the stock price shock since otherwise the Metropolis-Hasting procedure would identify a very small or even negative  $\sigma$ , which then allows all the fluctuations to

	Prior distribution			Posterior stock price shock			Posterior cost-push shock		
	Type	Mean	St. error	Mean	5%	95%	Mean	5%	95%
$\sigma_\pi$	Uniform	[0,1]	–	0.0012	0.0009	0.0014	0.0120	0.0104	0.0134
$\sigma_y$	Uniform	[0,1]	–	0.0032	0.0022	0.0045	0.0176	0.0159	0.0192
$\sigma_s$	Uniform	[0,1]	–	0.0421	0.0329	0.0512	–	–	–
$\sigma_x$	Uniform	[0,1]	–	–	–	–	0.1798	0.1622	0.1951
$\rho_s$	Beta	0.85	0.1	0.5408	0.4671	0.6129	–	–	–
$\rho_x$	Beta	0.85	0.1	–	–	–	0.6321	0.6266	0.6367
$\sigma$	Normal	1	0.2	0.8875	0.6205	1.1418	1.4460	1.4460	1.4460
$\gamma$	Normal	0.3	0.2	0.3155	-0.0131	0.6374	0.2998	-0.0394	0.6547
$\omega$	Beta	0.66	0.05	0.6563	0.5820	0.7315	0.6567	0.5769	0.7351
$\phi_\pi$	Normal	1.300	0.1	1.1508	1.0004	1.3160	1.5305	1.5305	1.5305
$\tilde{\nu}$	Normal	0.0	0.3	-0.0112	-0.5266	0.4722	-0.0059	-0.4961	0.5497
Log data density:				1326.97			653.54		

Table A.1: Priors and parameter estimates of the Bayesian estimation

be solely explained by demand shocks. Since this is not a realistic feature it is excluded by slightly decreasing the prior.

The model's key parameter is  $\nu$ , the marginal costs' elasticity to stock prices, which is measured to be very close to zero. It furthermore is not well identified with a standard deviation of the posterior of about 0.5, reflecting the value of its prior. This sheds further doubt on the usefulness of asset price targeting as a tool to stabilize economic activity since, at least for European data, it is unclear to which degree asset prices entail a risk for the economy.

As one would expect and as it is reflected by the higher log data density, exogenous fluctuations in asset prices provide a better data fit than the additional cost push shock. The cost push shock enters the model similarly (though not identically) to the productivity shock while with a stock market shock it is possible to completely steer the stock price dynamics exogenously. However, in the absent of endogenous fluctuations in asset prices such model is unable to capture the tail distribution of asset price movement with respect to interest rates.

Comparing these results to the results from the grid-based maximization technique, most parameter estimates are confirmed. However, although estimating  $\sigma_y$  to be slightly smaller than for the model with endogenous fluctuations, the degree of exogenous fluctuations in stock prices needs



to be considerably high while being in particular unable to match the correlation between output and inflation (model with stock price shock: -0.5718).

Unfortunately, a strong exogenous component in stock price fluctuations lacks economic intuition and empirical evidence. With exogenous fluctuations in asset prices the vast majority of the asset price dynamics are due to the stock price shock. This does not link well to any known story, including news shocks, and lacks economic intuition. While news shocks could explain exogenous movements to some extent, the fact that stock prices appear to be driven by highly persistent news shocks in combination with interest rate setting does not seem very convincing.

### A.3 Solving for the rational expectations equilibrium

In expectations it has to hold that

$$E_t \tilde{v}_{t+1}^\pi = \rho_\pi \tilde{v}_t^\pi \quad (\text{A.3.1})$$

$$E_t \tilde{v}_{t+1}^y = \rho_y \tilde{v}_t^y. \quad (\text{A.3.2})$$

Using this form, the PLM can be written using the system of equations (2.3.1) – (2.3.5) and bring all expectations to the LHS:

$$\beta E_t \pi_{t+1} = \pi_t - \kappa x_t - \tilde{v}_t^\pi \quad (\text{A.3.3})$$

$$E_t y_{t+1} = \sigma^{-1} r_{t+1} + y_t - \tilde{v}_t^y \quad (\text{A.3.4})$$

$$\beta E_t s_{t+1} = s_t + r_{t+1} \quad (\text{A.3.5})$$

$$0 = -x_t + \eta y_t + \dot{i}_t - \nu s_t \quad (\text{A.3.6})$$

$$0 = -\dot{i}_t + \phi_\pi \pi_t + \phi_s s_t \quad (\text{A.3.7})$$

$$E_t \tilde{v}_{t+1}^\pi = \rho_\pi \tilde{v}_t^\pi \quad (\text{A.3.8})$$

$$E_t \tilde{v}_{t+1}^y = \rho_y \tilde{v}_t^y \quad (\text{A.3.9})$$

Using (A.3.6) and (A.3.7) to substitute out  $i_t$  and  $x_t$  and rewriting as a matrix yields the System (2.4.2). Let me for rewrite this system as

$$\tilde{\mathbf{P}}E_t\tilde{\mathbf{x}}_{t+1} = \mathbf{M}\tilde{\mathbf{x}}_t$$

and  $\tilde{\mathbf{N}} = \tilde{\mathbf{M}}^{-1}\tilde{\mathbf{P}}$  is the  $5 \times 5$  matrix which summarizes the dynamics of the perceived law of motion of rational agents. I use eigenvector/eigenvalue decomposition to obtain  $\mathbf{\Gamma}\mathbf{\Lambda}\mathbf{\Gamma}^{-1} = \tilde{\mathbf{N}}^{-1}$ , where  $\mathbf{\Lambda}$  is the diagonal matrix  $\text{diag}(\lambda_1, \lambda_2, \dots, \lambda_5)$  of the eigenvalues of  $\tilde{\mathbf{N}}^{-1}$  ordered by size (smallest in modulus first) and  $\mathbf{\Gamma}$  the associated eigenvectors, columns ordered in the same fashion. The expectation system can then be rewritten as

$$\mathbf{\Gamma}^{-1}E_t\mathbf{x}_{t+1} = \mathbf{\Lambda}\mathbf{\Gamma}^{-1}\mathbf{x}_t.$$

Let me denote the sub-matrix of  $\mathbf{\Lambda}$  that only contains unstable eigenvalues as  $\mathbf{\Lambda}_u$ , and the associated eigenvectors I will likewise call  $\mathbf{\Gamma}_u^{-1}$ . I then know that  $\mathbf{\Gamma}_u^{-1}E_t\mathbf{x}_{t+1} = \mathbf{0}$  if I want to be consistent with transversality or feasibility constraints. Using this fact I can solve for  $E_t\mathbf{x}_{t+1}$ :

$$E_t\mathbf{x}_{t+1} = \mathbf{\Gamma}_{u,1:3}^{-1}\mathbf{\Gamma}_{u,4:5}E_t\tilde{\mathbf{v}}_{t+1} = \mathbf{\Gamma}_{u,1:3}^{-1}\mathbf{\Gamma}_{u,4:5}\rho\tilde{\mathbf{v}}_t$$

Note that the requirement that  $\mathbf{\Gamma}_{u,1:3}$  is invertible implies the Kuhn-Tucker condition: I impose that  $\mathbf{\Gamma}_{u,1:3}$  is a square matrix with full rank. This means that the number of forward looking variables has to equal the number of unstable eigenvalues  $\lambda > 1$  of  $\tilde{\mathbf{N}}^{-1}$ . Accordingly, I do only look at policy parameters  $\phi$  that guarantee that this condition is met. Let me define  $\bar{\mathbf{\Omega}} = \mathbf{\Gamma}_{u,1:3}^{-1}\mathbf{\Gamma}_{u,4:5}$ . The solution from the main body is then  $\bar{\mathbf{\Omega}}_{1:2,1:2}$ .<sup>2</sup>

---

<sup>2</sup>This implies that rational agents do not take asset prices into account when forming expectations. However, a more general approach including the adjustment for measurement errors of projecting three endogenous variables on two shock terms (stochastic indeterminacy) approximately lead to the same  $\Omega$ . Assuming that agents use OLS to regress  $\mathbf{x}_t$  on  $\tilde{\mathbf{v}}_t$ ,  $\tilde{\mathbf{\Omega}} = (\bar{\mathbf{\Omega}}^T\bar{\mathbf{\Omega}})^{-1}\bar{\mathbf{\Omega}}^T \in \mathbb{R}^{3 \times 2}$  and  $\tilde{\mathbf{\Omega}}_{1:2,1:2} \approx \bar{\mathbf{\Omega}}$ .

## A.4 Eigenvalues of the Rational Expectations System

In (2.3.6) I define the matrix  $\mathbf{N}$ , which represents the dynamics under Rational Expectations and is crucial for the dynamics under quasi-rational expectations. For both cases the Blanchard-Kahn condition for a determined and unique solution of the rational expectations is that the eigenvalues  $\lambda$  of  $\mathbf{N}^{-1}$  are lying outside the unit circle. Although an analytical representation exists, I opt for the numeric representation for reasons of clearer presentability. Figure A.1 shows the eigenvalues depending on both Taylor-Rule parameters and key parameter  $\nu$ , the elasticity of external finance premium with respect to equity. The dashed red line marks the point in parameter space where the relevant eigenvalues turn negative.

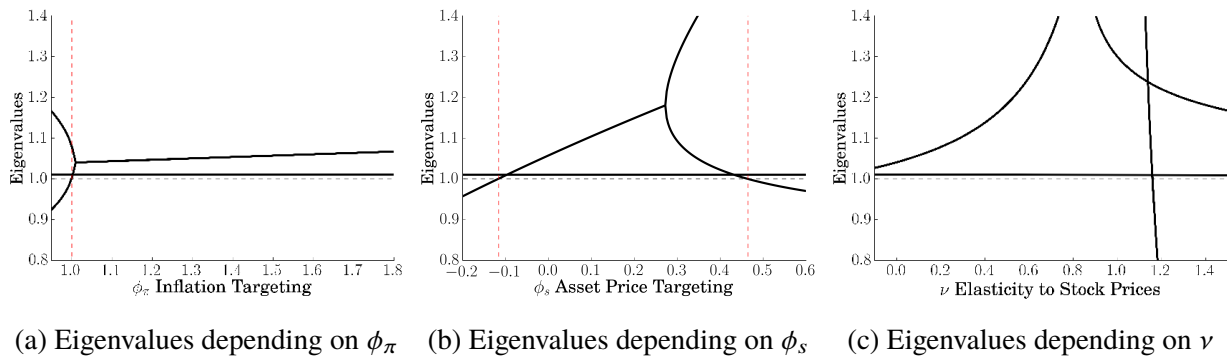


Figure A.1: Eigenvalues of the underlying model ( $\mathbf{N}^{-1}$ ) as a function of monetary policy parameters and the elasticity of the external finance premium.

## A.5 Fitting the model to match the statistical moments from Table 1

Let me denote the  $2 \times 3$  matrix that summarizes the standard deviations and cross correlations between the endogenous variable as  $\Xi$ . I run simulations of 20,000 periods and denote the same statistical moments of a simulation by  $\tilde{\Xi} : (\nu, \sigma, \Delta) \rightarrow \mathbb{R}^{2 \times 3}$ . This matrix is a function of  $\nu$ ,  $\sigma$  – the standard deviation of both shock terms – and  $\Delta$  which includes the behavioral parameters  $\alpha$  and  $\gamma$ .

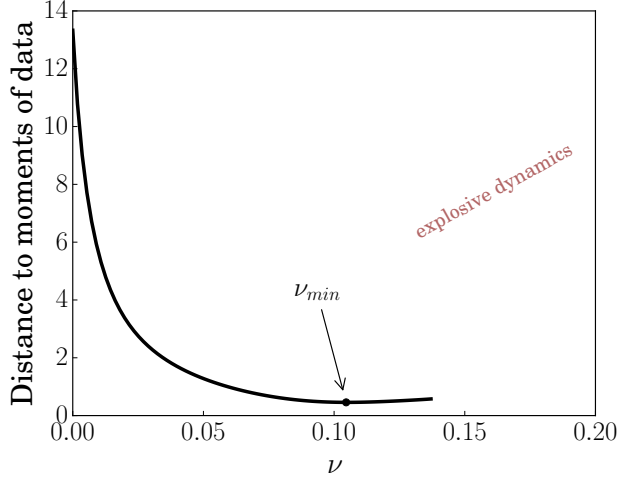


Figure A.2: Norm of the Matrix of squared errors. Here as a function of the marginal cost elasticity to stock prices

As the key fitness measure I use the Frobenius norm  $\chi = \|\Xi/\tilde{\Xi} - 1\|$  of the percentage deviation of both matrices.

Due to the nonlinearity and particularly the non-continuity it is not feasible to use standard optimization algorithms. The simplest solution is to use brute force in combination with a grid. I construct a grid of 5000 points for each parameter value. Each iteration, I only take one parameter (let me use  $\psi \in \{\nu, \sigma, \Delta\}$  as the placeholder for a parameter) and calculate a vector  $\chi_\psi$  containing the  $\chi$  for each point on the grid. At the end of the iteration, set  $\psi = \arg \min \bar{\chi}_\psi$  and repeat the process with the next parameter (i.e. select a new  $\psi$ ). Cycle through the parameter space until the parameter values converge. Here the convergence criterion  $\frac{\psi_i - \psi_{i-1}}{\psi_i} < 0.001$  is chosen, where  $i$  indicates the number of the iteration.

This will, with finite precision, approximate a local minimum of the non-continuous, non-smooth mapping. This implies sensitivity to the initial parameter values of the procedure. To ensure the global minimum I created a sparse grid of initial parameter values. This is in particular important because most numerical minimization procedures will halt when NaNs are encountered.

Figure A.2 plots the fitness measure as a function of  $\nu$ . While for low values the distance is very high, fit increases considerably for values larger  $\approx 0.07$ . For values larger than  $\approx 0.14$  dynamics are explosive and no measure is available.

# Chapter 3

## On the Evolutionary Fitness of Rationality

### 3.1 Introduction

*“If any group of traders was consistently better than average in forecasting stock prices, they would accumulate wealth and give their forecasts greater and greater weight. In this process, they would bring the present price closer to the true value.”*

— Cootner, 1964

The above quote encapsulates one of the cornerstones of economic theory: the Rational Expectations Hypothesis holds that in an efficient market, agents that do not act rationally will be outperformed by those who move wisely and well-informed. Agents that underperform for a longer period are then driven out of the market, leaving only rational agents behind (Friedman, 1953). But it also postulates that rational agents will foresee the fundamental price, instead of the *true* price, which might be affected by sentiment traders. *Perfect Rationality* however implies not only the knowledge on the economic fundamentals, but also on the market itself and its participants.

To study the interaction of perfectly rational and sentiment traders in a market with positive feedback and to reassess the Rational Expectations Hypothesis I extend the Brock-Hommes Heterogeneous Agent Switching Model (Brock and Hommes, 1998, referred to as BH) to incorporate fully rational agents. These agents are *hyper-rational* as they do not only know the market environment,

but also they are aware of the behavior of any other agent in the market, their expectations on the future, their relative number today and, in expectations, for all periods to come. Put differently, they are rational in the classical sense and they are aware of the presence of non-rational agents – their expectations are model consistent and do not contain any misspecification.

Using the BH98 model this work focuses on two questions. First, I investigate whether rational agents are actually stabilizing the market. This problem is particularly interesting because there are two contradicting intuitions. On the one hand, the presence of rational agents could indeed have a stabilizing effect. These agents foresee other agents behavior perfectly and might be able to outsmart them. Technically, the rational expectations solution is a fixed point in the law of motion: the expected value of the price next period determines the outcome at the present period while it is also path-dependent upon the latter. On the other hand, the presence of rational agents is likely to be destabilizing. Since rational agents know all other beliefs, they anticipate the behavior of boundedly rational agents, trade accordingly and thereby positively reinforce the behavior and beliefs of bounded rational agents. In this case the Rational Expectation Hypothesis could never hold, because sentiment beliefs are ex-ante true through their reinforcement by rational agents. Secondly and much related, I ask whether a reasonably large fraction of boundedly rational agents can survive in the medium term and long term.

The concept of Rational Expectations has first been introduced by Muth (1961) and in particular gained popularity through the work of Lucas (1972). Although the leading paradigm today, it also received early critique as most prominently by Simon (1955), and counter-critique e.g. by (Sims, 1980, “the wilderness of bounded rationality”). Kirman (1992) is a more contemporary response to the *rational expectations revolution*, which has redrawn attention on the matter. As of yet, this debate is unsettled. Although many major economists agree that the assumptions underlying the Rational Expectations framework might be too demanding, as of yet a clear and commonly accepted alternative has not emerged. For a deeper discussion on the concept of bounded rationality see f.i. Conlisk (1996).

With closer ties to the scope of my work here, the idea of “a superior analyst” that is capable of outsmarting less sophisticated agents has also been discussed in the seminal article of Fama (1965). In his view a sufficiently large fraction of smart agents would be able to prevent bubbles. In contrast, De Long et al. (1990) support the idea that rational traders might amplify price swings that are induced by noise traders, if rational agents can anticipate other agents trading behavior and can act ahead of time. Other than in their model, I explicitly consider the behavior of boundedly rational agents and allow for the fraction of rational agents to be endogenous. Amir et al. (2005), Hens and Schenk-Hoppé (2005) present evidence that non-CAPM trading strategies might under some circumstances be the only evolutionary stable strategy, whereas Evstigneev et al. (2002) provide evidence that such strategies might be able to consume the whole market. Work of Blume and Easley (2010) investigate similar questions in a different setup. They empathise the fact that the market selection hypothesis fails when markets are incomplete and discount factors are heterogeneous.

The state of the art of the research conducted on bounded rationality in the BH-tradition can be found in Hommes (2013). Using not only behavioral models but also laboratory experiments, this branch of research emphasises that the stability of a system crucially depends on the type and degree of expectations feedback. Negative feedback loops, although not under every circumstances, tend to be relatively stable since expectations are not self-enforcing. In the case of a positive feedback loop, stability depends crucially on the magnitude of the eigenvalues. In this work I focus on financial markets since in this area the positive feedback is apparent. I do not discuss the case of commodity markets with negative feedback here since this problem normally embeds an explicit dynamic system also under rational expectations and is well known to the literature as the so-called *hog-cycle model* (Brock and Hommes, 1997).

Fully rational agents in a financial market context have, in this stand of research, only been studied in Brock et al. (2009). The authors focus on the question whether market completeness can improve price stability while also touching on the question of whether rational agents stabilize the market. Rationality is then derived from the perfect foresight argument while ruling out bubble solutions, i.e. reducing the degrees of freedom of the perfect foresight solution by one. Instead

of finding an explicit representation of the implied actual law of motion, they provide analytical results and conclude that whether rational agents can stabilize the market or not highly depends on the degree and composition of boundedly rational fraction of agents in the market.

The key methodological contribution of this work is an iterative numerical method to solve for an explicit representation of the rational expectations equilibrium. To my best knowledge, such methods have not been applied to the form of highly nonlinear models with endogenous fluctuations as the one considered here. Methods of this type are well known to dynamic economic theory and the field of recursive macroeconomics (e.g. Ljungqvist and Sargent, 2012). See Judd (1998) and Miranda and Fackler (2004) for comprehensive, general surveys of numerical methods in economics. The iterative method considered here allows to explicitly account for fully rational agents in an Heterogeneous Agent Switching Model (HAM), whereas the previous literature focussed on concepts that embed a closed form solution for the law of motion. As such, the model was mainly studied with fundamentalist traders or approximating rational agents by using the concept of the perfect foresight path. These concepts do not require to solve for the rational expectations solution, but, as I explain further below may not conceptualize the rational expectations solution correctly or, as for fundamentalists, not be well suited to answer the question that drives this work.

The rest of this work is structured as follows. In Section 3.2 I briefly introduce the model and sketch the numerical solution method in Section 3.3. In Section 3.4 I present and discuss the simulation results. Section 3.5 concludes.

## **3.2 Model**

This section attempts to stay as close as possible to the original model of Brock and Hommes (1998) as this model is well established in the literature of bounded rationality and nonlinear economic dynamics and provides a well-known reference point. I hence follow their derivation of the model closely.



Accordingly, let us consider a stylized asset market with a continuum of agents that are neither constrained in borrowing nor in short-selling. Furthermore, each trader is a myopic mean variance maximizer, which implies that trader  $i$ 's demand  $z_{i,t}$  for the risky asset is a linear function of his beliefs  $x_{i,t+1}^e$  about the price in  $t + 1$  as well as today's price. If  $x_t$  is defined to be the percentage deviation of the price for the financial asset at time  $t$  from its fundamental value, market clearing reads then as a no-arbitrage condition of the form

$$Rx_t = \int x_{i,t+1}^e di,$$

where  $R$  stands for the (time-invariant) discount rate. This equation will also be called the *law of motion* (LOM) of the model.

For simplicity this work is restricted to a family of models with a maximum of three types of agents of which two types are symmetrical. This is sufficient for the purpose of this work and covers a wide range of possible dynamics while still allowing for a parsimonious model. The solution concept for the Rational Expectations path could however easily be adapted to more complicated models with more types.

In particular agents are either sentiment traders, i.e. optimistic or pessimistic about the near future, or perfectly rational and then  $x_{i,t+1}^e \in \{x_{t+1}^{e+}, x_{t+1}^{e-}, E_t x_{t+1}\}$ . The beliefs  $E_t x_{t+1}$  of this third group of perfectly rational agents are formed based on the information available at  $t$  and the complete knowledge of the model. I present and discuss the solution concept for the expectations of these agents in the next section.<sup>1</sup> Let me formalize the predictors of sentiment traders by

$$x_{t+1}^{e+} = +\beta \quad \text{and} \quad x_{t+1}^{e-} = -\beta,$$

---

<sup>1</sup> In the rest of this work I am using the terms *rational* and *rational expectations* interchangeably. To be precise, agents could form perfectly rational expectations but act boundedly rational or even irrationally [with respect to utility/profit maximisation]. In models of the BH-type, in fact all agents *act* fully rationally *given their beliefs*. Their choice of predictors however might not be completely rational. This, again, is to remain consistent with the majority of the literature.

where the degree of sentiment bias is denoted by  $\beta$ . Market clearing is then given by

$$Rx_t = (1 - n_{+,t} - n_{-,t})E_t x_{t+1} + (n_{+,t} - n_{-,t})\beta, \quad (3.2.1)$$

where  $n_{+,t}$  denotes the fraction of optimists and, likewise,  $n_{-,t}$  the fraction of pessimists. Further following Brock and Hommes (1998), these fractions are updated according to the *performance measure*  $\pi_{i,t}$  for each predictor. As such, *realized profits* is a natural candidate and I define:

$$\pi_{i,t} = (x_t - Rx_{t-1})(x_{i,t}^e - Rx_{t-1}) - \mathbb{1}_{RE}\kappa, \quad (3.2.2)$$

where  $\kappa$  is the cost for obtaining the rational expectations solution and  $\mathbb{1}_{RE}$  an indicator function that equals one if the agent is rational. All other predictors are costless. The first part of the first term at the RHS of (3.2.2) denotes the actual resale value of the asset minus the opportunity costs for financing the purchase in the previous period. The second part represents agent *i*'s demand, resulting from the mean-variance maximization given the agent's past belief about the price. The choice of the performance measure is an essential ingredient of the model. It determines the properties of the dynamic system. Realized profits from trading qualify in several ways for this model. Instead of receiving a high pay-off for an accurate estimate of the price, in order to receive a positive profit it is sufficient to have made a correct choice on whether to go short or long. Likewise, a trader *A* that has a strong positive belief about next periods prices will invest more money in the asset than another trader *B* with a relatively lower positive forecast. Even if *B*'s forecast was perfectly correct, trader *A* will still make higher profits since he invested more. This feature, i.e. that profits are non-proportional to forecasting errors, is unique to financial markets and captured by Equation (3.2.2).

The probability that an agent is of type  $i \in \{+, -, RE\}$  is determined by a *multinomial discrete choice model* depending on the past performance of the predictor:

$$n_{i,t} = \frac{e^{\delta\pi_{i,t-1}}}{\sum_{j \in N} e^{\delta\pi_{j,t-1}}}. \quad (3.2.3)$$

If a predictor is relatively more successful than others, it is more likely to be chosen, hence the fraction of agents using this predictor increases.  $\delta$  is called the *intensity of choice* which governs the speed of switching between predictors. If  $\delta \rightarrow \infty$ , all agents will immediately switch to the most successful predictor. This completes the full specification of the formal model.

### 3.3 Numerical solution

We are looking for a representation such that at any point  $t$  the state of the system can be calculated given only the past states. The model presented in the previous section does not allow for such a solution in closed form. In the literature on nonlinear economic dynamics it is sometimes argued that, in the absence of stochastic disturbances, the so called Perfect Foresight Path (PFP) coincides with the Rational Expectation Path (REP), where the former can then be used to simulate the model. As I show here, this does not hold in general. In particular, the PFP embeds one additional degree of freedom and includes the REP as a special case. To clarify, let us provide two definitions.

Consider a dynamic forward looking model  $h$  that represents the state at time  $t$  by

$$y_t = h((y_{t+1})^e, y_{t-1}, \dots, y_{t-k}), \quad (3.3.1)$$

where  $h$  is some mapping from  $\mathbb{R}^{k+1}$  to  $\mathbb{R}$  and  $(\cdot)^e$  is an expectation operator that is yet to be defined.

**Definition 1** (Rational Expectation Path). *Given a sufficiently long history  $\{y_{t-1}, y_{t-2}, \dots\}$ , the Rational Expectation Path (REP)  $\{y_t\}_k^\infty$  satisfies (3.3.1) in each period  $t > k$  and  $E_t y_{t+1}$  equals the expected value of  $y_{t+1} = h(\cdot, y_t, \dots)$  based on the information set implied by the history at time  $t$ .*

*The Rational Expectations solution associated with the REP can then said to be **ex ante model consistent**.*

Put differently,  $E_t y_{t+1} = E[y_{t+1} | y_t, \dots, y_{t-k}]$ , and each  $y_t$  on the REP with  $t \in \mathbb{R}^+$  must be consistent with  $E_t y_{t+1}$ . Hence,  $E_t y_{t+1}$  represents a fixed point. In fact, in the absence of stochastic shocks, it is deterministic and coincides with the actual  $y_{t+1}$ . Note that this implies that each  $y_t$

on the REP is a mapping  $y_t : \mathbb{R}^k \rightarrow \mathbb{R}$ , meaning that only the history of length  $k$  is necessary to compute  $y_t$ .

**Definition 2** (Perfect Foresight Path). *Given a dynamic forward looking model  $h$  and a sufficiently long history, a time series each  $z_t$  on the Perfect Foresight Path (PFP)  $\{z_t\}_0^\infty$  satisfies if it satisfies*

$$z_t = \{y_{t+1} : y_t = h(y_{t+1}, y_{t-1}, \dots, y_{t-k}) | y_t, y_{t-1}, \dots, y_{t-k}\},$$

*i.e. it is imposed that  $\{y_t\} = z_{t-1}$  and every  $y_{t+1}$  is chosen given  $y_t$ . The Perfect Foresight solution associated with the PFP can then said to be **ex post model consistent**.*

Let us emphasise that the PFP implies that  $z_t$  is a mapping  $z_t : \mathbb{R}^{k+1} \rightarrow \mathbb{R}$ . Now a history of length  $k + 1$  is necessary to find  $z_t$  which is a larger set than what actually occurs in  $h$ . In Appendix B.1 I summarize the intuition behind the following proposition:

**Proposition 1.** *The Perfect Foresight Path (PFP) coincides with the Rational Expectations Path (REP) only if all initial conditions of the PFP lie on the REP.*

*Proof.* See e.g. Blanchard and Kahn (1980). □

It is easy to see that this equivalence condition is not satisfied in general. Since it can only be guaranteed once the REP is known, the Perfect Foresight Solution is not a particularly helpful tool when solving dynamic models with rational expectations, independently of whether they are deterministic or stochastic.

The numerical procedure used here to solve for the REP is closely related to the method of *Fixed Point Iteration* (see e.g. Judd, 1998) which is well known to the literature of e.g. dynamic games or nonlinear macroeconomics. The basic intuition goes as follows: at any point in time  $t + s > t$  the future is unknown and hence can not be used to find the systems' state at  $t$ . If we however would have a solution to the system, it could be used as well to infer on any future state  $x_{t+s}$ . The existence of such representation implies a solution for REP. The method first assumes existence of a solution and then verifies by finding its exact representation. This is in fact equivalent to a rational agent

facing a decision that will affect the future, while the future is relevant to make the decision - the so-called *fixed point argument* that is implied by  $E_t y_{t+1}$  being a function of  $y_t$ .

Plugging (3.2.2) in (3.2.3) and inserting the result into (3.2.1), the model's state at  $t$  can be expressed as a (known, nonlinear) function  $f$  that depends on the rational expectation of next periods price, as well on the past values:

$$f : (E_t x_{t+1}, x_{t-1}, x_{t-2}) \rightarrow \mathbb{R}.$$

Note that in the absence of shocks it also holds that  $E_t x_{t+1} = x_{t+1}$ , which however is not a necessary condition for this method to work. Being able to solve for the Rational Expectations Path of this system implies that there exists a recursive representation  $g$  that is a (unknown) function of only the history in  $f$ :<sup>2</sup>

$$g : (x_{t-1}, x_{t-2}) \rightarrow \mathbb{R}. \tag{3.3.2}$$

$g$  can be found by inserting it into  $f$ . We know:

$$\begin{aligned} x_t &= f(x_{t+1}, x_{t-1}, x_{t-2}) \\ &= g(x_{t-1}, x_{t-2}) \end{aligned}$$

and

$$\begin{aligned} x_{t+1} &= g(x_t, x_{t-1}) \\ &= g(g(x_{t-1}, x_{t-2}), x_{t-1}). \end{aligned}$$

Then our problem boils down to finding a function  $g$  that satisfies

$$g(x_{t-1}, x_{t-2}) = f(g(g(x_{t-1}, x_{t-2}), x_{t-1}), x_{t-1}, x_{t-2}), \tag{3.3.3}$$

---

<sup>2</sup>Note that in a linear framework this would imply the Transversality Condition since it is a sufficient condition for the existence of a recursive solution. This can be seen e.g. by using Eigenvalue-Eigenvector Decomposition. Numerical procedures of this type generally do not require the Transversality Condition for stationarity.

which can be done numerically. For this purpose, let us define  $\mathbf{x} = \{X_1, X_2, \dots, X_M\}$  a vector of  $M$  grid points on which  $g$  shall be defined.  $g$  then resides on  $\mathbb{R}^{M \times M}$ , which in the two-dimensional case is a matrix. Given an initial guess  $g_0$  we can iterate (3.3.3)

$$g_{k+1}(\mathbf{x}, \mathbf{x}') = f(g_k(g_k(\mathbf{x}, \mathbf{x}'), \mathbf{x}), \mathbf{x}, \mathbf{x}').$$

If  $g$  exists,  $\|g_{k+1} - g_k\|$  converges to zero when  $k$  goes to infinity. The iteration halts once a  $\|g_{k+1} - g_k\| < \epsilon$  for some predefined very small  $\epsilon$  is reached.<sup>3</sup>

### 3.4 Results

In this section five different types of experiments are presented. First, I explore the potential dynamics of the system by varying the behavioral parameters  $\delta$  and  $\beta$ . Second, to study the effect of rational traders in the market I compare the dynamics of different values of costs for rational expectations  $\kappa$ . Then I compare these results with a model including fundamentalists instead of rational agents. I furthermore revisit the BH two-trader type model. Lastly, to identify further mechanisms and to provide robustness I compare these results to a model with a risk-adjusted fitness measure.

$R$	$\delta$	$\beta$	$\kappa$
$0.99^{-1}$	1	1	0

Table 3.1: Benchmark parametrisation

As the benchmark the parameters given in Table 3.1 are used. The two behavioral parameters  $\delta$  and  $\beta$  are normalized to unity for simplicity and costs for rational expectations are set to zero.

---

<sup>3</sup>In the Appendix I deepen further on conditions under which the iteration might not converge.

### 3.4.1 Endogenous dynamics

In order to assess the general dynamics I am looking at the dynamics under different values for the behavioral parameters  $\beta$  and  $\delta$ . This also sheds light on the question whether boundedly rational agents might get driven out of the market, although the discrete choice model does not allow for a zero-fraction of sentiment traders. This however can also be seen as a realistic feature since in real markets there will always be new market entrants that might as well be boundedly rational. The same strategy is chosen by Brock et al. (2009).

In Figure 3.1 the long-run dynamics of prices and fractions of each type of agents are shown as a function of the intensity of choice  $\delta$ . For low values of this parameter, the simulations suggest that the steady state is stable and unique. As  $\delta$  increases, a limit cycle emerges after what appears to be a Hopf-Bifurcation at  $\delta \approx 1.2$ , which is a typical characteristic of a 3-trader-type-model. The amplitude of these cycles increases in  $\delta$ .

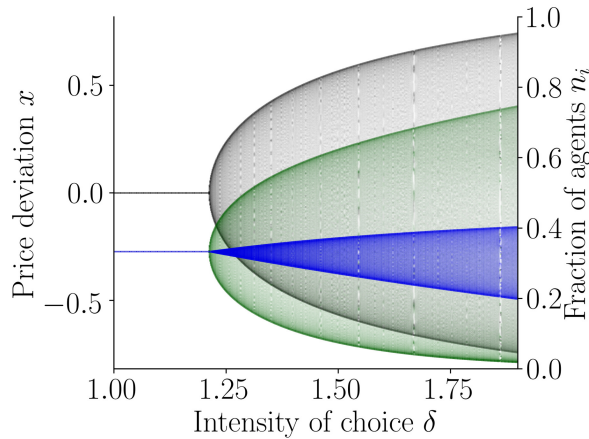


Figure 3.1: Bifurcations w.r.t.  $\delta$ . In blue (green) and on the right axes the dynamics of the fraction of rational (sentiment) traders.

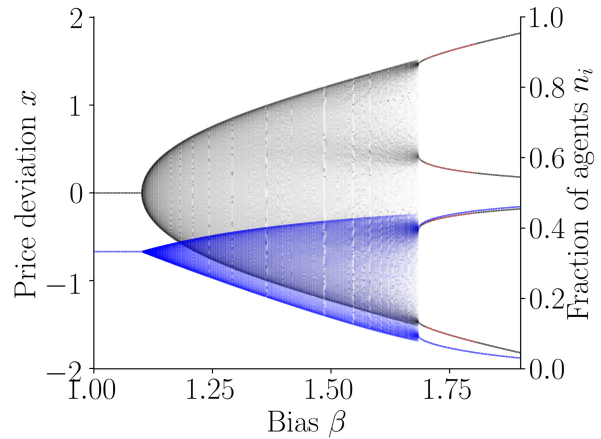


Figure 3.2: Bifurcations w.r.t.  $\beta$ . In blue and on the right axes the dynamics of the fraction of rational traders.

An increase in the intensity of choice does also lead to an increase in the variation of the amount of rational agents, that is depicted in the blue region, while the green region shows fluctuations in the amount of optimistic agents.<sup>4</sup> Why does the number of sentiment agents vary significantly

<sup>4</sup>Note that, due to the symmetric construction of optimist and pessimist traders, the dynamics of optimist and pessimist traders are in fact symmetric.

stronger than the number of rational agents? In a bearish period the number of optimist traders will fall because the belief that prices are high is not profitable. Hence, in this period most boundedly rational agents will be pessimists. While rational agents always forecast the price correct, their profit and hence their selection only depends on  $x_t - x_{t-1}$ . Pessimist traders however will gain higher profits since they have overestimated the change in price and hence decreased their short position more strongly. This relatively higher profit, in extreme periods, does also motivate some of the rational agents to turn pessimistic. Because the intensity of choice governs how strongly agents react to incentives by higher profits, the overall fluctuations in optimist/pessimist traders increases with  $\delta$ . The dynamics for  $\delta \rightarrow \infty$  can be seen in Figure 3.5, which turns out to be a 4-cycle similar to the original BH-model.

Figure 3.2 shows the dynamics with respect to the degree of optimism and pessimism  $\beta$ . While these dynamics are qualitatively similar to those in Figure 3.1, the amplitude increases faster. The more biased agents are, the stronger they influence the price in troughs and peaks. Up from a value of  $\beta \approx 1.7$  the system converges as well to a stable 4-cycle. This can be seen as further evidence against the aforementioned hypothesis: starting from a small fraction of boundedly rational agents, if their beliefs are strong enough they are again able to influence the market in such a way that rational agents will partially adopt and reinforce their beliefs in their forecast.

This already reveals, given the model assumptions here, that the hypothesis that boundedly rational agents are driven out of the market in the long run might be controvertible. If agents tend to switch to the more successful strategy more quickly, there is always a fraction of boundedly rational agents that is sufficiently successful to gather at least a share of the market. Any rational agent then adjusts his belief accordingly and by doing so amplifies the boundedly rational traders' beliefs. Although this model does not explicitly allow for agent types to be driven out of the market, the fact that the agents' fraction fluctuates around one-third suggests a falsification of the REH.



### 3.4.2 The consequences of rationality

Let us now have a look at Figures 3.3 and 3.4 where the costs for the rational expectations operator are plotted against the x-axis. This allows to study the effect rational agents have, given that an increase in costs is associated with a decrease of the number of rational agents.

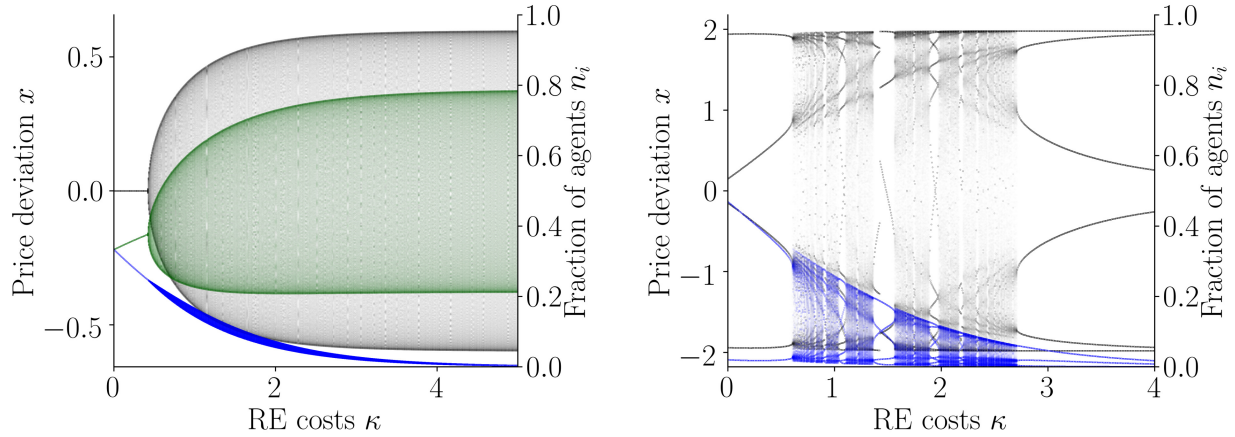


Figure 3.3: Bifurcations w.r.t.  $\kappa$ .  $\beta = 1.05$ . Figure 3.4: Bifurcations w.r.t.  $\kappa$ .  $\beta = 2$ . In blue Blue and green are the fractions of rational and positive biased agents. dynamics of the fractions of rational agents.

The effect of an increase in costs for rationality is twofold. As suggested by Figure 3.2, the steady state for  $\kappa = 0$  is stable and unique for  $\beta = 1.05$ . Taking this as a starting point, in Figure 3.3 I let  $\beta = 1.05$  and increase  $\kappa$ . The blue area confirms that the fraction of rational agents decreases with the costs for rationality while simultaneously the fraction of sentiment traders increases. Once this fraction is large enough, again a Hopf-Bifurcation occurs and endogenous speculative dynamics arise. A further decrease in the number of rational agents does not seem to have large impact on the price dynamics, and in fact the dynamics will be identical for any value of  $\kappa > 4$ . In this setup the presence of a significant number of rational agents does add stability and inhibits or at least mitigates the degree of endogenous fluctuations. This experiment hence supports the hypothesis that a reasonably large fraction of rational agents can indeed stabilize the market and bring prices closer to fundamentals.

The implications of Figure 3.4 however draw a rather ambiguous picture. Here sentiment traders are biased more strongly which, as suggested by Figure 3.2, leads to a 4-cycle even if rationality is

costless. Dynamics then become quasi-periodic for values of  $\kappa$  smaller than approximately 2.7 are suggested by the diagram for higher  $\kappa$ . A further decrease in the number of rational agents then leads to a stable 6-cycle as  $\kappa$  goes to infinity. This result can be interpreted such, that once – for one reason or another – endogenous fluctuations arise, a variation in the fraction of rational agents can not add further stability (and costs can not be set  $< 0$ ). This is well in line with the previous argument that either a high intensity of choice or stronger biases after some point affect the market such, that rational agents to some extent have to adapt the boundedly rational beliefs.

### **3.4.3 Comparing rational and fundamentalist traders**

Fundamentalists traders are agents who believe that the price will always return to its fundamental value, which is here given by 0. In a narrow sense they are also rational since in the absence of sentiment traders, the zero steady state would be the rational solution.

Comparing those with rational agents, the dynamics are less stable with rational agents than with fundamentalists instead. This can be seen in Figure 3.5: a lower intensity of choice is required to offset the cyclic behavior induced by biased traders. This suggests that the hypothesis of an amplifying moment of rational agents is true as well: rational agents anticipate the behavior of boundedly rational agents and their resulting trading behavior induces a further moment of destabilization. This result is intuitive since fundamentalists will always believe that the future price equals its fundamental value. Their belief will hence will not be affected by the fact that that (other) boundedly rational traders have stronger beliefs or belief-switching occurs faster. This is in particular important for values of  $\delta$  (here between 1.2 and 1.6) or  $\beta$  where rational agents already take the behavior of sentiment traders into account and hereby amplify their impact on the market.

### **3.4.4 The two-trader type model**

Let us now turn to the 2-trader type model with rational agents vs. trend followers. In this model, there are only two types of agents, i.e. fully rational and trend chasing traders. The beliefs of trend

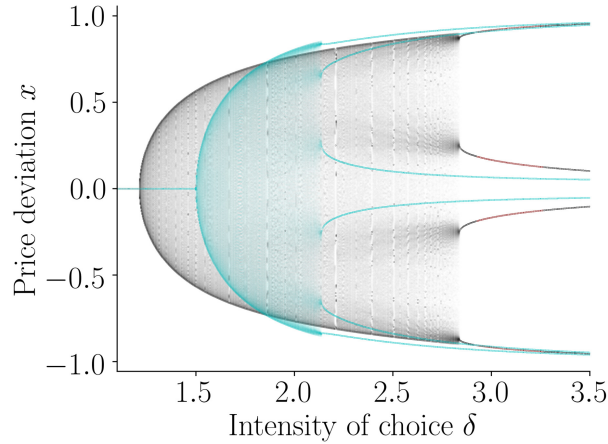


Figure 3.5: Bifurcations w.r.t.  $\delta$ . In cyan the same simulation with fundamentalists instead of chasers is given by

$$x_{t+1}^e = \gamma x_{t-1},$$

where  $\gamma$  denotes the degree of trend extrapolation. Note that, since the past prediction is part of the profit equation, the numerical method is now a mapping from  $\mathbb{R}^3 \rightarrow \mathbb{R}$  which has consequences for the calculation speed.<sup>5</sup>

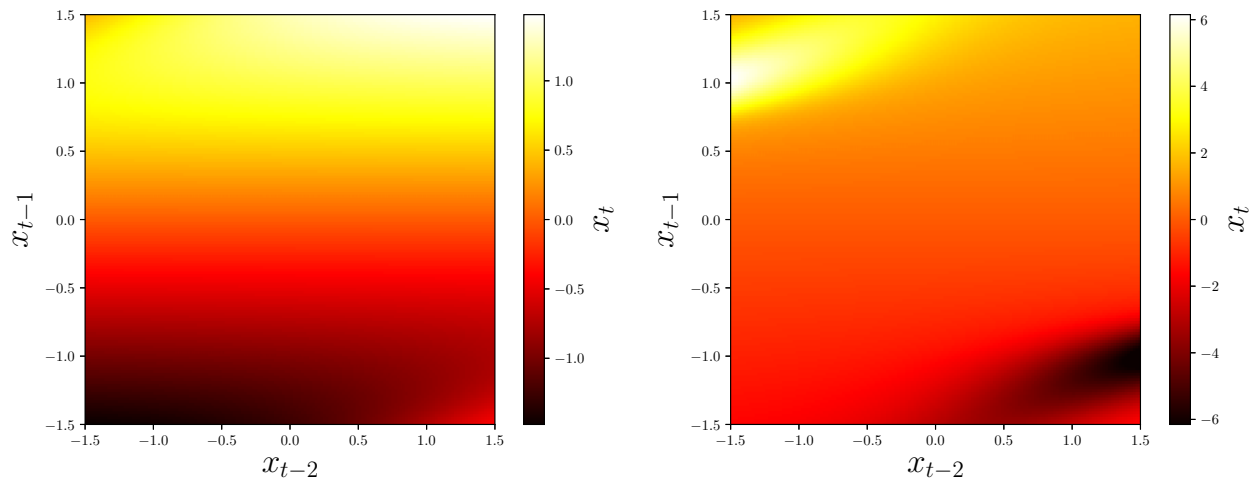


Figure 3.6: Heatmap of  $g_{75}(x_{t-1}, x_{t-2}, 0)$  for  $\gamma = 0.2$  (l) and  $\gamma = 0.25$  (r). For the lower gamma, the zero steady state is still locally stable. For  $\gamma = 0.25$  trajectories diverge slowly outwards. Note the different scales.

<sup>5</sup>For  $\gamma \neq 0$  the function in (3.3.2) is defined on 3-D space, i.e.  $g(x_{t-1}, x_{t-2}, x_{t-3})$  due to an extra time lag in the fractions  $n_{i,t}$ .

The actual dynamics are considerably simple. Given the parameters in Table 3.1 for  $\gamma$  smaller than approximately 1.24, the zero steady state is unique and stable. As  $\gamma$  increases further, the zero steady state becomes unstable and the system's dynamics explode. This phenomena can be described by a so-called *hard bifurcation*.<sup>6</sup> The underlying forces are represented in Figures 3.6 and 3.7. Note that the functions represented here are actually two dimensional slices of the 3-dimensional function  $g$ . The fraction of (boundedly) rational agents is, after transition, constant and equal 0.5 for all values of  $\gamma$ . This is explained by the fact that in the zero steady state, the expectations of both types of agents are perfectly correct and switching probabilities are equal. When the system explodes, initialized by a positive price the trajectory is dominated by the beliefs of the trend chasers and, since agents that are forming rational expectations are always right, payoffs again are equal, leading to equal fractions of agent types.

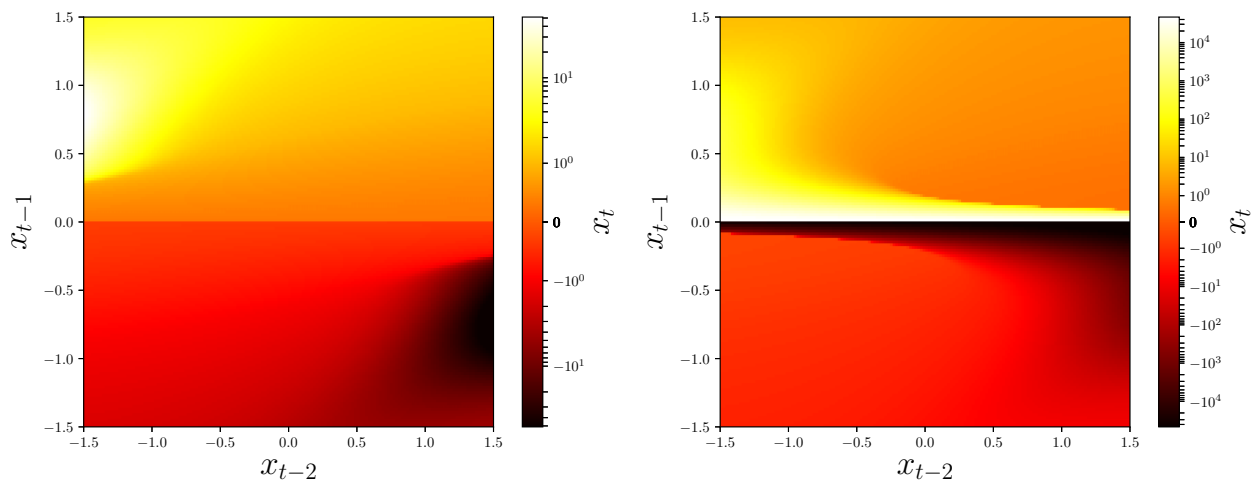


Figure 3.7: Heatmap of  $g_{75}(x_{t-1}, x_{t-2}, 0)$  for  $\gamma = 0.32$  (l) and  $\gamma = 0.4$  (r). Trajectories lead from the center to the periphery. This effect amplifies with higher  $\gamma$ . Note the different scales.

This setup is difficult to asses and compare to the original results of Brock and Hommes (1998) and does not appear to link to their Lemmas 2 – 4, where the dynamics with respect to  $\gamma$  are related to values of  $2R$  or  $R^2$  respectively. While in their setup chaotic dynamics could be identified, this is

<sup>6</sup>To allow for comparison with the original results of Brock and Hommes (1998) I also conducted this simulation with their calibration of  $R = 1.1$ . The hard bifurcation is then, taking potential measurement errors into account, at  $\gamma \approx 1.363$ .

not the case for the model described here.<sup>7</sup> It however strikes intuitive that, as soon as the degree of trend extrapolation reaches a certain threshold (which is well in line with their analytical results), rational agents will start to follow the trend chasers beliefs. This then further amplifies their beliefs, without any remaining force to revert the trajectory back to the steady state.

### 3.4.5 Risk adjusted fitness measure

As pointed out above, the fact that boundedly rational agents constitute a notable fraction of agents is due to the fact that in certain phases of the speculative cycle, these agents make higher profits than rational agents. This in turn is true because profits, and with that payoffs and feedback, are not proportional to agents' forecasting errors. Again, as empathised in Brock and Hommes (1998), this can be attributed to the fact that the profit equation in (3.2.2) is correctly specified, but does not take investment risk into account. In particular, the payoff function considered before does adjust for the variance of asset prices. For this reason payoffs that are proportional to intra-period forecasting errors are another natural candidate for this analysis.

Hommes (2013, p.166 f.) shows that the payoff function then takes the form<sup>8</sup>

$$\pi_{i,t} = -(x_t - x_{i,t}^e)^2 - \mathbb{1}_{RE}\kappa. \quad (3.4.1)$$

This rather reads as a punishment for forecasting errors than a payoff. It is immediately clear that the payoff for rational agents is *always*  $-\kappa$ , while payoffs for sentiment traders are  $-(x_t \pm \beta)^2$ . Then it is apparent that this system may have three steady states, each of them being associated to the

---

<sup>7</sup>Note that proving chaos is in general nontrivial. Likewise it is hard to show the absence of chaos for all variation of parameter values, in particular since simple results like *period-3 implied chaos* in general do not hold for the multidimensional case.

<sup>8</sup>This takes into account the assumption of mean-variance utility and adjusts for the potential risk-free profits. In fact the full payoff function is then given by

$$\pi_{i,t} = -\frac{1}{2a\sigma^2}(y_t - y_t^e + \epsilon_{y,t})^2,$$

where  $a$  stands for the agents' risk-aversion and  $\epsilon_{y,t}$  the stochastic fluctuations in prices with standard deviation  $\sigma$ . Since in 3.2.3 this term will be multiplied by  $\delta$  for which no empirical counterpart is available, a precise adjustment for  $a$  and  $\sigma^2$  would not provide further insight. Since I am focussing on the endogenous fluctuations in this work, the noise term can also be omitted.

dominance of one agent type. Figure 3.8 shows the associated bifurcation diagrams where steady states exchange stability at a Pitchfork bifurcation. Here the steady state where pessimistic traders dominate is ruled out because the simulation is initialized with a positive price. Note that for any  $\kappa > 0$  a further increase of  $\kappa$  has the same effect on rational agents as an increase in  $\delta$  since the fraction of rational agents evolves proportional to  $e^{-\delta\kappa}$ , which is relevant in particular since the fundamental steady state is stable and unique for all  $\delta > 4.1$ . This implies that, given a high intensity of choice, the costs of being rational have to be relatively small to shift the system back to a non-fundamental steady state.

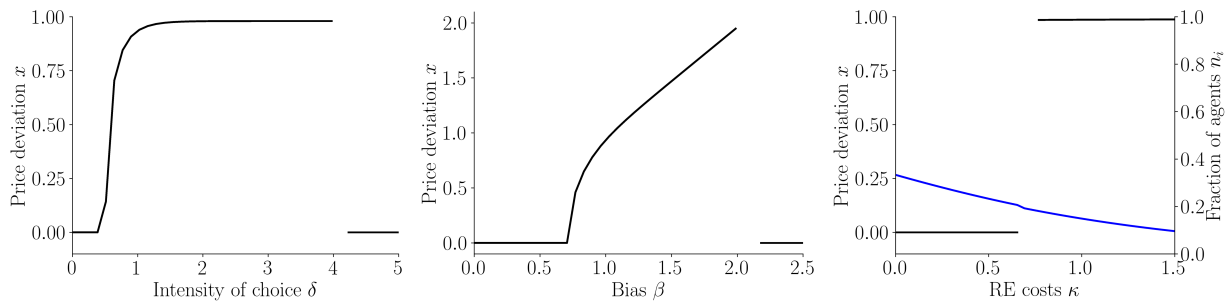


Figure 3.8: Bifurcation diagrams for  $\delta$  (left),  $\beta$  (center) and  $\kappa$  (right). For the first two parameters the steady state increases with the parameter but falls back after a certain threshold. The last Figure depicts the case where  $\delta = 5$ . The blue line shows the fraction of rational agents.

These results again confirm the hypothesis that rational agents are stabilizing, in particular in favor of the fix-point argument. Note that we can not observe endogenous fluctuations in any of the simulations. Rather, a fix point is identified which is then either in accordance with the belief of one of the sentiment traders, or – if intensity of choice or bias is sufficiently strong – returns to the fundamental value where sentiment traders’ beliefs cancel each other out. The results for the simulations with increasing  $\kappa$  show clearly that a decreasing fraction of rational agents then leaves the market to either of the sentiment traders.

## 3.5 Conclusion

This work studies the dynamics of a simple financial market that is characterized by the coexistence of perfectly rational agents and sentiment traders, while the total number of each type varies proportional to their market performance. To find a solution to such model, one of my central contributions is to make use of iterative methods to find the rational expectations path.

The primary finding is that rational agents are prone to adapt beliefs of boundedly rational agents, which might amplify endogenous trading dynamics. This result is mainly driven by the self-fulfilling nature of asset price expectations, and amplified if fitness measures do not account for risk. This lack of stabilization by perfectly rational traders stems for the fact that they anticipate the behavior of boundedly rational agents and use this information in their trading decision.

Secondly, the numerical evidence sheds doubt on the proposition that boundedly rational agents are driven out of the market in the long run. Although with certain limitations, this is due to the strong feedback of expectations on prices and the fact that net profits are not proportional to forecasting errors. Depending on the magnitude of individual beliefs, speculative dynamics can emerge and coordination of rational and boundedly rational traders can become complicated, however not chaotic.

Further, in most setups the presence of rational agents does indeed tend to stabilize the market. This is in particular true if agents' fitness measure accounts for taken risk. However, the stable price might not necessarily reflect the economic fundamentals. Decreasing the amount of rational agents tends to destabilize the market, a result which again depends on the strength of sentiment beliefs. When beliefs are moderate, decreasing costs and increasing the degree of rationality in the market might in fact facilitate coordination and stabilize prices. If individual biases are strong the net-stabilizing effect might however be negligible.

In conclusion, my findings here support the hypothesis that rational agents tend to *stabilize* the market, but sentiment traders are not in general driven out of the market and can have considerable impact on prices. Given payoffs that do not account for risk correctly, the presence of boundedly rational agents also might induce endogenous oscillations that are further amplified by rational

agents. Fully rational agents are then “riding the wave” and behave *as if* they are boundedly rational. These results shed further doubt on the propositions that financial markets are stable due to Rational Expectations Hypothesis.



# Appendix B

## B.1 Intuition behind Proposition 1

To illustrate, let us take the most basic example of a forward looking dynamic system of the form

$$y_t = \alpha E_t y_{t+1}, \quad |\alpha| < 1.$$

If we iterate this forward to  $E_t y_{t+1} = \alpha E_t y_{t+2}$  and repeat this step infinitely many times the rational expectations solution is

$$y_t = \lim_{s \rightarrow \infty} \alpha^s E_t y_{t+s} = \begin{cases} 0 & \forall \lim_{s \rightarrow \infty} E_t y_{t+s} \in \mathbb{R} \\ \{\} & \text{if } |\lim_{s \rightarrow \infty} E_t y_{t+s}| = \infty. \end{cases}$$

This result is intuitive since  $\lim_{s \rightarrow \infty} \alpha^s = 0$  for all  $|\alpha| < 1$ . If then the absolute value of  $\lim_{s \rightarrow \infty} E_t y_{t+s}$  is infinite,  $y_t$  equals the product of zero and infinity which has no solution. But the above implies that any expectation  $E_t y_{t+k}$  given  $k > 0$  can either be 0 or has no solution, but an infinite solution is ruled out. It then follows that the Rational expectations solution, if it exists, is

$$y_t = 0 \quad \forall t.$$

This reveals a common misconception. The Transversality condition

$$\lim_{s \rightarrow \infty} E_t y_s = 0$$

guarantees a solution, but does not directly imply stationarity. This example can be generalized to the multidimensional case and has been treated rigorously in the literature, see f.i. Blanchard and Kahn (1980) for the conditions on existence of a solution expressed in terms of the eigenvalues of the respective system.

The general argument for the PFP implies  $E_t y_{t+1} = y_{t+1}$  and that at time  $t - 1$  the value of  $y_t$  must have been known in order to be able solve for  $y_{t-1}$ . By assumption, this value must have also been correct since agents have perfect foresight. It follows directly that it can be solved for  $y_t$  by just iterating the law of motion back one period. So if  $y_t = \alpha y_{t+1}$ , then it must also hold that  $y_t = \alpha^{-1} y_{t-1}$  because  $y_{t-1}$  has already been chosen in the anticipation of  $y_t$ .

This conclusion however is false, which can be illustrated by looking at the stability characteristics of both systems under  $|\alpha| < 1$ . As shown above the RE system is stable and jumps back to zero from every point in  $\mathbb{R}$ . The perfect foresight path diverges unless the initial value of  $y_{t-1}$  lies on the REP i.e. is equal to zero. The system diverges for any set of initial values that does not lie on the RE path. Divergence then implies that the initial value for  $y_{t-1}$  must have already been off the REP, which is a direct contradiction to the conjecture that every perfect foresight solution also satisfies the condition of rationality.

Also, if agents possess complete knowledge of the system's LOM and the states that are relevant for this LOM, there is no reason to impose that further past information should be necessary to solve the expectations problem (such as  $y_{t-1}$  in the case of the example here).

## B.2 A note on convergence

While for the examples outlined here my algorithm (almost)<sup>1</sup> always converges, convergence is not guaranteed for any *finite* grid  $\mathbf{x}$ . If we continue Figure A.3.6 with higher values of  $\delta > 6$ , even after many iterations the solution generally does not satisfy the convergence criterion. This problem is caused by the discontinuity of  $g$  for  $\delta \rightarrow \infty$ . To understand, imagining the numerical function  $g$

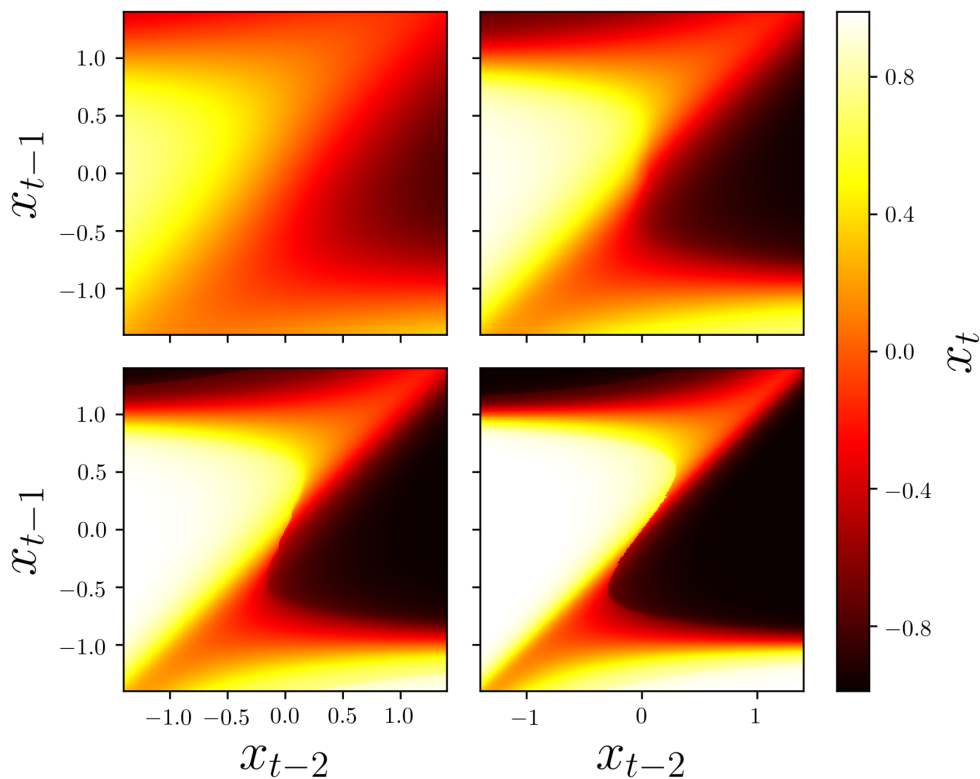


Figure B.1: Illustration of the numerical function  $g$  for  $\delta = 1.1$  (ul),  $\delta = 2.5$  (ur),  $\delta = 4$  (ll) and  $\delta = 7$  (lr). When  $\delta$  increases the function becomes steeper and at the diagonal a small change in  $x_{t-1}$ ,  $x_{t-2}$  leads to a larger change in  $x_t = g(x_{t-1}, x_{t-2})$ .

on the grid  $(\mathbf{x}, \mathbf{x}')$ . In my solution algorithm it is necessary to evaluate  $g_k(g_k(\cdot), \mathbf{x})$ , i.e. to have a

<sup>1</sup> In Figure A.3.7 the region around  $\beta \in (1.7, 1.8)$  does not converge. The same occurs for  $\delta$  between approximately 3 and 3.4 in Figure 3.5. I use red to mark the values where the procedure did not converge.

real valued input into function defined on a discrete space, which numerically necessarily involves an interpolation. Let us assume we want to evaluate  $g_k$  at a point  $z \in \mathbb{R}$  for which no  $g_k(z, \cdot)$  exists, let us call  $X_k$  the nearest smaller grid point  $X_j < z$  for which  $g_k(X_j, \cdot)$  exists, and  $X_{j+1} > z$  the nearest higher point respectively. Accuracy of the interpolation then naturally decreases with  $\Delta G = |g_k(X_j, \cdot) - g_k(X_{j+1}, \cdot)|$ .

Figure B.1 shows  $g_{500}$  for increasing values of  $\delta$ . While for  $\delta = 1.1$  the function is very smooth, it becomes steeper when  $\delta$  gets larger. For  $\delta = 7$  in the last diagram the function is very steep with  $\Delta G$  being particularly high in the center and decreasing in the periphery. This region close to the center and along the diagonal is not captured well by the algorithm since small deviations in the respective  $g_n(\cdot)$ -values lead to large differences in  $g_{n+1}(\cdot)$ -values in the next iteration. This problem can be partially mitigated by increasing the size of the grid and inserting a new node  $X_{\text{new}}$  such that  $X_k < X_{\text{new}} < X_{k+1}$ . Now  $g(z, \cdot)$  will be evaluated as the interpolation between  $X_k$  and  $X_{\text{new}}$  (assuming  $z < X_{\text{new}}$ ) which probably exhibits a lower  $\Delta G$  (though not with certainty).

In fact, for every finite grid there will always be a combination of parameters for which there exists a  $X_k$  and  $X_{k+1}$  for which  $\Delta G$  is large. This problem can generally be tackled quite efficiently by implementing an Endogenous Grid Method that allocates relatively more grid points to the critical region. This however is not necessary for the example here. Since we know that the center will reflect the trajectory back to the periphery, and by noting that the periphery also redirects to the periphery we can conjecture that the center has little effect on the simulation of the time series. For this reason, if necessary, I take the 500th iteration of  $g$ ,  $g_{500}(\cdot)$ , and use it to simulate the time series.<sup>2</sup> The result confirms that the conjecture was correct.

---

<sup>2</sup>In fact in most cases  $g_{25}$  is already sufficient and accurate.

# Chapter 4

## Capital Taxation and Investment: Matching 100 Years of Wealth Inequality Dynamics

### 4.1 Introduction

*“All animals are equal*

*But some animals are more equal than others”*

— George Orwell, *Animal Farm* (1946)

In November 2016 investor George Soros lost about one billion US\$ when betting against the market during the market race that greeted the election of Donald Trump to become the 45th president of the United States.<sup>1</sup> Meanwhile, others like the billionaire Warren Buffett have put billions into stock markets, betting on the increase in government investment and the accompanying effects on stock prices. Thus, the disagreement in expectations about the evolution of the asset market has a major impact on the overall wealth of individuals and also influences the distribution of wealth. Given the fact that the top wealth holders are heavily invested, even small disagreements will have large impacts and finally shape the distribution of wealth.

---

<sup>1</sup> As stated by news agencies such as *CNBC*:  
<http://www.cnbc.com/2017/02/17/george-soros-loaded-up-with-big-bets-against-the-stock-market.html>

In this work we develop a formal model that explicitly incorporates the role of disagreement in expectations and shows its ability to match the dynamics of the upper tail of the empirical distribution of wealth. Our second key contribution is not only to match the correct levels of wealth, but we are also able to replicate the transition dynamics for four countries in the long run – the US, UK, Sweden, and France – by only using time series of income and capital gains taxes as a model input.

The interest in the distribution of wealth gained momentum after the publication of Piketty (2014) in which he documents an increasing concentration of wealth starting in the 1980s. For the important case of the United Kingdom and the USA, Piketty (2014) argues that the tax policy under Margaret Thatcher respectively Ronald Reagan was one major factor leading to the increase in wealth inequality. Yet, the book does not feature a formal model to discuss the claim. Using the novel data evidence, both Kaymak and Poschke (2016) and Hubmer et al. (2016) employ extended Bewley type models in order to explain the observed wealth inequality dynamics and, in particular, the fat tails in the wealth distribution. They confirm that the evolution of wealth inequality can to a large extent be explained by changes in the tax system. In particular, Hubmer et al. (2016) argue that other potential mechanisms such as the witnessed increased in income inequality (the key mechanism generating wealth inequality in Bewley models), the falling labor share, or the  $r > g$  story of Piketty (2014) all fall short of accounting for the data. We argue that these models are highly complex and require numerical solutions and thus might obscure the fundamental underlying mechanism.

The model employed in this work can be classified as being a *random growth* model in the tradition of Benhabib et al. (2011) which has been recently calibrated to US data in Aoki and Nirei (2017). In our work, we present a closed-form solution of both the stationary distribution and the transition dynamics as emphasized in Gabaix et al. (2016). The key underlying mechanism in these type of models is idiosyncratic investment risk creating the fat tails in the distribution of wealth. As formally shown in Benhabib et al. (2015) – in which the idiosyncratic investment risk is introduced into an otherwise standard Bewley economy – the investment risk is important to

explain the top-wealth inequality. In general, Bewley models inherit the wealth inequality from the income inequality. As it is well-known from the empirical evidence the top inequality of the stock measure of wealth is far higher than for the flow measure of income making Bewley models incapable of producing top wealth inequality.<sup>2</sup> Usually, it is argued that the idiosyncratic shocks can be generated by idiosyncratic asset holdings. This work extends the literature by also including expectation disagreement. We show that wealth (approximately) follows a symmetric double Pareto distribution in line with the empirical evidence of Benhabib et al. (2014). Without expectation disagreement the distribution would be characterized by a fat tail only in the right part of the distribution (the *rich*) and proper calibration to also match the transition speed would overestimate idiosyncratic risk.

We show furthermore that taxation is a key ingredient to understanding wealth inequality. In particular, without taxation inequality there would be no finite inequality.<sup>3</sup> Note that we focus on the top-shares - i.e. the right tail of the wealth distribution. It is well-known that the low share at the low end of the Lorenz curve can be explained by different forms of (financial) frictions - most prominently borrowing constraints.<sup>4</sup> In this work, we abstract from these frictions. Moreover, by assuming constant relative risk aversion individuals have identical risk exposure and identical marginal propensities to consume. In such a world, given that traders share identical beliefs, there would be no trade. Instead, we let them make random and considerably small forecasting errors that are normally distributed around the Rational Expectations solution. Thus - and in contrast to the partial equilibrium model employed in Benhabib et al. (2011) - this work features a closed economy.

---

<sup>2</sup>Other extensions proposed in the literature to match the top wealth share problem is the consideration of entrepreneurs in the tradition of Quadrini (2000) or Cagetti and Nardi (2009a) having different savings patterns than the rest of the population.

<sup>3</sup>A similar point was made in Fernholz and Fernholz (2014). This work, also featuring a slightly different model, only relied on simulations, while we can provide a closed-form solution.

<sup>4</sup>Other factors are different access to financial markets (entry barriers for the poor, lower transaction costs for the wealthy) or heterogeneous access to important information and expertise. We abstract from these systematic frictions in favor of the wealthy. Sometimes the role of portfolio structuring (in particular a utility function with decreasing relative risk implying higher return portfolios for wealthy), different consumption patterns (heterogeneous marginal propensity to consume) and wealth transmission are emphasized.

With this micro-founded - yet simple - model we are able to capture both the level as well as the dynamics<sup>5</sup> of wealth inequality for the USA, UK, France, and Sweden for a very long time period ranging up to 100 years.<sup>6</sup> To our best knowledge we are the first to present a comprehensive cross-country view of inequality as the existing literature has focused solely on the USA (also being the country with the highest increase in wealth inequality among the covered countries). The different levels of taxation are key to understanding the differences in top wealth shares across countries and in time.

The remainder of this work is structured as follows. In Section 4.2 we provide an overview of the empirical and theoretical literature on wealth inequality with a focus on recent papers trying to fit the empirical evidence. In the following section, we present the micro-foundations for our formal model and discuss analytic statistic properties in Section 4.4. Section 4.5 compares these analytical results with the empirical findings. Section 4.6 uses the model to generate forecasts about the future evolution of wealth inequality and presents robustness checks. Finally, Section 4.7 wraps up.

## 4.2 Literature

Following the major public debate surrounding the publication of the work of Piketty (2014), the interest in distributional measures has recently increased. In particular, the empirical evidence regarding inequality - especially for the flow measure of income - has substantially improved. Cross-country evidence is assembled and made freely available on the *World Income & Wealth Database* maintained by the collaborative effort of many researchers.<sup>7</sup> Despite this effort the data availability of consistent and long-run measures of wealth inequality is still highly limited. The database provides long-run data for both the United States of America and the United Kingdom.

---

<sup>5</sup> As argued in Gabaix et al. (2016) the problem with this type of model is that the convergence dynamics are usually very slow.

<sup>6</sup> The first three countries are the only three countries with long-run data for wealth inequality, as documented in the *World Wealth & Income Database*. Sweden is covered extensively in the work of Lundberg and Waldenström (2017).

<sup>7</sup> The data is available at [wid.world](http://wid.world).



The US data was updated recently by Saez and Zucman (2016). The latest data update for the UK was conducted by Alvaredo et al. (2017). The quality of the French data (especially from the 1970s onward) was recently substantially improved by Garbinti et al. (2017). Evidence for Sweden is compiled by Daniel Waldenström and his collaborators (Lundberg and Waldenström, 2017).<sup>8</sup> A recent comprehensive survey on the overall empirical evidence regarding wealth inequality is given in Roine and Waldenström (2015). The discussion about the distribution of wealth is not an end in itself, but also contains important policy implications as it impacts on the conduct of monetary policy (Kaplan et al., 2016) and also influences economic growth (Clemens and Heinemann, 2015).

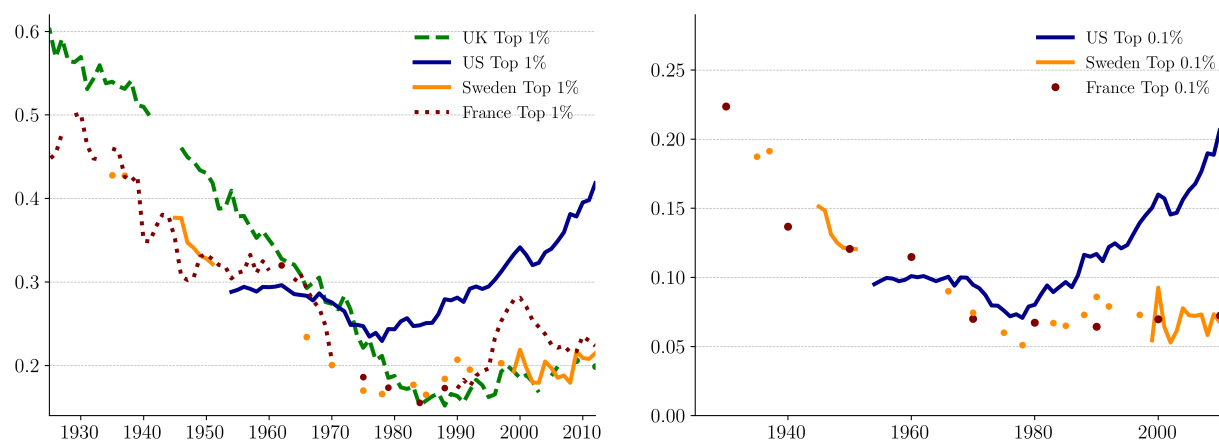


Figure 4.1: Top wealth shares. Data source: `wid.world` and Lundberg and Waldenström (2017) for Sweden.

Figure 4.1 present evidence on the top-shares for the Anglo-Saxon countries - the USA and the UK - and the European Countries - France and Sweden - in the long run.<sup>9</sup> While there is an overall decrease in all countries until the 1980s, inequality has subsequently increased. This increase is modest in the European countries, but highly pronounced in the USA. In particular, it emerges for the top wealth holders. In general, inequality is higher in the USA. As we will discuss in the course of this work this can be traced back to the different taxation systems.

Different theoretical models compete in order to explain the witnessed degree of inequality. Usually, models in the Bewley-type tradition are considered in order to discuss inequality (Bewley,

<sup>8</sup> The data is freely available on his homepage.

<sup>9</sup> Note that there is no top 0.1% data available for the United Kingdom. Thus, the graph only displays the USA, France and Sweden.

1977; Huggett, 1993; Aiyagari, 1994). Yet, it has been formally shown by Benhabib et al. (2011) that these types of models - build around the notion of additive idiosyncratic labor income risk - will fail to generate the fat tails in the wealth distribution and thus match the shares of the top wealth holders. Benhabib et al. (2011) propose a model with multiplicative idiosyncratic capital income risk in order to replicate the current state of wealth inequality in the USA. They follow an argument laid out as early as Wold and Whittle (1957), building on random growth. Thus, this type of literature is often referred to as *random growth models* (Gabaix et al., 2016). Taxation of capital (income) plays a crucial role in these models. Using simulations, Fernholz and Fernholz (2014) show that wealth inequality does explode without redistribution in a standard model with idiosyncratic investment risk. We build a similar model, yet introduce dispersion of individual opinion about the prospects of an investment in order to match the shares of the top wealthy individuals.

While heterogeneous portfolios are often motivated by different degrees of risk aversion (and marginal propensities to consume), this argument does not hold for the very rich. In such models, the highly wealthy should have a relatively similar portfolio structure. This, however, does not take into account the vast variety of different asset classes and almost infinite supply of similar assets within classes. Moreover, due to heterogeneous individual expectations about future prospects, agents will hold different positions in identical assets. For that reason we motivate heterogeneous portfolios by marginal disagreement on future returns, of which a considerable degree is documented by Greenwood and Shleifer (2014b) in a survey over six data sets on investor expectations of future stock market returns. Furthermore, evidence from the lab has shown that individuals generally do not to a good job when forming expectations and these expectations are furthermore largely heterogeneous.<sup>10</sup> If agents are heterogeneous in their projections, they also invest in different assets.

Thus, compared to the literature in the tradition of Benhabib et al. (2011) assuming some exogenous random noise in a partial equilibrium setting, our model is closed in a general equilibrium tradition by allowing individuals to trade with each other. Moreover - and compared to for example

---

<sup>10</sup> For a recent overview on the heterogeneous expectations hypothesis see for example Hommes (2013).

Fernholz and Fernholz (2014) - our work goes beyond pure numerical simulation and is able to not only quantify the stationary distribution but the whole dynamics of the top wealth shares. Similar to the existing literature we are able to match the steady state of wealth inequality.

The second aim of this work is to take the available data and test whether the model can match in the transition. This is of particular importance as it was recently argued in Gabaix et al. (2016) that while these models capture the steady state of inequality well, the dynamics in these models are far too slow compared to empirical evidence. Both closed form solution and simulations confirm that our model also matches the dynamics of inequality well.

Of course, using the new data evidence, similar projects have been undertaken. Most prominently, Kaymak and Poschke (2016) use the evidence for the United States from 1960 to the most present date to present a calibrated model in the Bewley tradition. Using the modification of Castaneda et al. (2003), allowing for extreme *superstar* shocks producing high levels of income inequality, the authors are able to match the data. With a very detailed modeling of the US-tax system (including income, corporate, and estate taxes as well as the pension system) the authors identify the contributing factors. They argue that the (exogenous) increase in income inequality and - for the distribution of wealth - the structure of the taxation and transfer system are highly important in order to explain the evolution of inequality.

A more comparable approach to this work is presented in Aoki and Nirei (2017), featuring a rich model in continuous time. In line with empirical evidence and due to idiosyncratic firm shocks the distribution of heterogeneous firms is given by Zipf's law.<sup>11</sup> The firm's income translates into income for private households, implying a realistic distribution of both income and wealth for private households. Combining this with tax rates, they are able to match both the dynamics and the state of inequality in the USA from the 1970s to the most recent years. Note that we focus on the distributional impact on taxation and do not make a statement about the macroeconomic impact of the wealth tax or even its optimal level.<sup>12</sup>

---

<sup>11</sup> The latter is a Power-law with an exponent  $\alpha = 1$ .

<sup>12</sup> As shown in the seminal contribution of Judd (1985) in a standard model the optimal tax on the stock value of wealth is zero. Yet, it is well-known that in Bewley-type models this result fails to hold and optimal taxes are positive in order to counteract excessive savings (Aiyagari, 1995). More recent contributions discussing the welfare impact in

Most closely connected to this work, Hubmer et al. (2016) extend an otherwise standard Bewley-type model with heterogeneous rates of time preference  $\beta_i$  in the tradition of Krusell and Smith (1998), with Pareto tails in the income distribution and idiosyncratic investment risk following Benhabib et al. (2011). They are able to quantitatively reproduce wealth inequality dynamics for the USA from the 1970s. As outlined before, they conclude that the most important mechanism driving the increases in wealth inequality is the change in the taxation system since the 1980s which is also central in our work. Compared to the existing literature focusing solely on the USA, we compare different countries and also use longer data evidence starting before World War II.

### 4.3 Model

We assume an economy with a large number  $n$  of individuals indexed by  $i$ . Their only income consists of investment returns and they are free to choose between a risk-free asset paying constant gross return  $R$  and a continuum of ex-ante identical risky assets of which each pays an idiosyncratic, stochastic dividend  $d_{i,t}$  every period  $t$ . To maximize their intertemporal consumption over an infinite time horizon the agents accumulate wealth  $w_{i,t}$ . Hence, each agent  $i$  faces the question of which amount  $c_{i,t}$  to consume and which amount  $x_{i,t} = z_{i,t}w_{i,t}$  of the risky asset to purchase. In this case  $z_{i,t}$  relates the demand for risky assets as a share of individual wealth  $w_{i,t}$ . Assuming log-preferences, the individual problem is then given by

$$\max_{c,z} \sum_{t=0}^{\infty} \beta^t \ln c_{i,t}$$

subject to the two constraints

$$c_{i,t} = (1 - s_{i,t})w_{i,t},$$

$$w_{i,t} = (R + [d_{i,t} + p_t - Rp_{t-1}]z_{i,t-1}) s_{i,t-1}(1 - \tau)w_{i,t-1}.$$

---

both the state and the transition of wealth taxes in the broader sense (including capital gain and inheritance taxes) are (among others) Castaneda et al. (2003), Domeij and Heathcote (2004), and Cagetti and Nardi (2009b).

Here we denote by  $\beta$  the intertemporal discount rate and  $s_{i,t}$  the savings rate. The value  $\tau$  captures a tax on the stock level of wealth.<sup>13</sup>  $p_t$  is the price for a risky asset in  $t$  and  $d_{i,t} = d + \epsilon_{i,t}^d$  its dividend with an idiosyncratic stochastic term  $\epsilon_{i,t}^d \sim N(0, \sigma_d)$ .<sup>14</sup> This problem does not directly entail a closed form solution, but can be separated into two stages that are relatively standard in the literature. Let us first solve the consumption problem.

Levhari and Srinivasan (1969) show that for log-utility, which is a particular case of Constant Relative Risk Aversion (CRRA) preferences<sup>15</sup>, in equilibrium agents consume  $1 - \beta$  of their wealth at the end of each period, i.e.  $s_{i,t} = \beta \forall i, t$ .<sup>16</sup> It is important to point out that this result holds despite the tax rate. Due to the exact offsetting of income and substitution effects for log-utility the savings rate is not distorted by the tax rate.<sup>17</sup> Note that the assumption of CRRA also explicitly avoids inequality dynamics induced by a different marginal propensity to consume. The fact that agents are not subject to borrowing constraints simplifies the model considerably, but is not realistic in the context of the lower 50% share of wealth holders. We hence focus on the upper tail of the wealth distribution. Then the law of motion for each individual's wealth follows

$$w_{i,t} = (1 - \tau)\beta \{R + (d_{i,t} + p_t - Rp_{t-1})z_{i,t-1}\} w_{i,t-1} \quad (4.3.1)$$

$$= (1 - \tau)\beta R_t^z(z_{t-1})w_{i,t-1}, \quad (4.3.2)$$

for which  $R_{i,t}^z(z_{i,t-1})$  summarizes the individuals gross return on investment.

---

<sup>13</sup> Note that in the empirical part it is very important that taxes vary in time. For the sake of readability we, however, suppress the time index in this section.

<sup>14</sup> Since those assets are ex-ante identical, their price is likewise the same.

<sup>15</sup> It is easy to show that for the special case  $\gamma \rightarrow 1$  the CRRA utility function boils down to log-utility  $\lim_{\gamma \rightarrow 1} \frac{c^{1-\gamma}-1}{1-\gamma} = \ln(c)$  for which income and substitution effects exactly cancel each other out.

<sup>16</sup> If for gross return  $1 + r$ ,  $r$  is given by a normal distribution with mean  $\mu$  and standard deviation  $\sigma$ , Levhari and Srinivasan (1969) show that the optimal value is given by  $s = (\beta \exp(\mu + 0.5\sigma^2)^{1-\gamma} \exp(-\gamma(1-\gamma)0.5\sigma^2))^{\frac{1}{\gamma}}$ . It is easy to see that for  $\gamma = 1$  we have  $s = \beta$ . Note that here we abstract from the taxes that would modify both  $\mu$  and  $\sigma$ . Given log-utility the latter, however, has no impact.

<sup>17</sup> The interested reader is also referred to Lansing (1999) showing that the seminal result of a zero optimal tax rate as proposed in Judd (1985) fails to hold with log-utility. For a more recent and general approach the reader is referred to Straub and Werning (2014).

For the second stage, in which we solve for the optimal demand for risky asset  $x_{i,t}$ , let us use Equation (4.3.1) to rewrite the maximization problem as

$$\max_z \sum_{t=0}^{\infty} \beta^t \ln\{(1 - \beta)w_{i,t}\} \quad \text{s.t.} \quad w_{i,t} = (1 - \tau)\beta R_{i,t}^z(z_{i,t-1})w_{i,t-1}$$

which is equivalent to

$$\max_z \sum_{t=0}^{\infty} \beta^t \ln \left\{ (1 - \beta)\beta w_{i,0} (1 - \tau)^t \prod_{k=0}^t R_{i,k}^z(z_{i,k-1}) \right\}.$$

Due to the logarithmic laws, the term  $\ln\{\prod_{k=0}^t R_{i,k}(z_{i,k-1})\}$  can be separated and is the only part that depends on  $z_t$ . Since we can rewrite  $\prod_{k=0}^t R_{i,k}^z(z_{i,k-1}) = R_{i,t}^z(z_{i,t-1}) \prod_{k=0}^{t-1} R_{i,k}^z(z_{i,k-1})$  this portfolio problem can be well-approximated by mean-variance maximization as laid out in Pulley (1983). The optimal demand for the risky asset  $x_{i,t}$  is then, up to a second order approximation, given by

$$x_{i,t} = z_{i,t} w_{i,t} = (E_t[d_{t+1} + p_{t+1}] - Rp_t) w_{i,t} / \sigma_d^2. \quad (4.3.3)$$

Note that – identical to the optimal consumption plan – the portfolio structure is independent of the wealth tax. As presented in Stiglitz (1969) for Constant Relative Risk Aversion preferences – of which the assumed log-utility is a special case – wealth taxation does not lead to a restructuring of the portfolio.

Market clearing requires  $\sum_{i \in n} x_{i,t} = X_t$ , with  $X_t$  being the total supply of the risky asset. Without loss of generality we can fix supply and normalize  $X_t$  to unity for all periods.<sup>18</sup>

We want to assume that return expectations are heterogeneous and each agent's expectation is a draw from the normal distribution around the rational expectation of future returns. Thus, the

---

<sup>18</sup> This implies that the market clearing price is defined by:

$$p_t = \{p : \sum_{i \in n} x_{i,t}(p, w_{i,t}) = 1\}.$$

Note that the optimal individual demand depends on individual wealth, which in turn also depends on the current price.

rational expectation operator  $E$  is replaced with a noise individual expectation operator  $\hat{E}_{i,t}$ , giving

$$\hat{E}_{i,t}[d_{t+1} + p_{t+1}] = d + E_t p_{t+1} + \epsilon_{i,t}^E, \quad \epsilon_{i,t}^E \sim N(0, \sigma_E).$$

Assume furthermore that no single person is rich enough or has an  $\epsilon_{i,t}^E$  large enough to influence the price.<sup>19</sup> Taking market clearing into account, it is clear that all wealth is owned by all agents and  $\sum_{i \in n} x_{i,t} = 1$ , i.e. aggregate demand for the risky asset equals one every period.

We want to assume that our taxation is redistributive transferring a lump-sum value  $T$  to all individuals  $T_t = \frac{\tau \sum_{i \in n} w_{i,t}}{n} = \frac{\tau W_t}{n}$  for an aggregate wealth  $W_t = \sum_{i \in n} w_{i,t}$ . Thus, no wealth is lost in the act of taxation. For our specific assumption, for which the aggregate wealth (being stationary) is identical to the number of agents  $W = n$ , this boils down to  $T_t = \tau_t$ .

Keeping this in mind and the aggregating over Equation (4.3.3) and (4.3.1) yields

$$p_t = R^{-1}(E_t p_{t+1} + d - \sigma_d^2 W_t^{-1}) \quad (4.3.4)$$

$$W_t = \beta(RW_{t-1} + d + p_t - Rp_{t-1}), \quad (4.3.5)$$

which is the law of motion for prices and aggregated wealth  $W_t$ . Since  $n$  is large, due to the law of large numbers idiosyncratic disturbances level out and aggregate wealth  $W_t = W$  is constant in the absence of aggregate shocks. The levels of prices and aggregated wealth thus reflect the detrended steady growth path.<sup>20</sup> Then we can also normalize the price to unity without explicitly accounting for market clearing. The steady state versions of (4.3.4) and (4.3.5) is

$$\sigma_d^2/W = d + 1 - R$$

$$W(\beta^{-1} - R) = d + 1 - R.$$

---

<sup>19</sup> This is indeed satisfied by the law of large numbers. The additional advantage of this assumption is, without loss of generality, that we can provide analytic results for the law of motion of individual wealth, aggregated wealth, and prices.

<sup>20</sup> This assumption implies that all growth in aggregate wealth can be attributed to increases in output, which constitutes a further simplification for the sake of simplicity. Since the key targets in this paper are wealth shared, these are unaffected by the overall level of wealth.

This implies that, given the normalization of prices,

$$W = \frac{\sigma_d}{\sqrt{\beta^{-1} - R}} \quad (4.3.6)$$

$$d + 1 - R = \sqrt{\beta^{-1} - R} \sigma_d. \quad (4.3.7)$$

Plugging Equation (4.3.3) into Equation (4.3.1), integrating individual forecast errors and setting prices to the steady state yields

$$w_{t,i} = \beta \left\{ R + \left( d + \epsilon_i^d + 1 - R \right) \left( d + \epsilon_{i,t}^E + 1 - R \right) \sigma_d^{-2} \right\} (1 - \tau) w_{t-1,i} + T_t.$$

Using Equation (4.3.7) and some algebra, the law of motion (LOM) for individual wealth can be written as

$$w_{i,t} = \beta \left\{ \beta^{-1} + \sqrt{\beta^{-1} - R} (\epsilon_{i,t}^d + \epsilon_{i,t}^E) \sigma_d^{-1} + \epsilon_{i,t}^d \epsilon_{i,t}^E \sigma_d^{-2} \right\} (1 - \tau) w_{i,t-1} + T_t.$$

For our simulations we use a quarterly calibration, so let  $\beta = 0.99$  and the real interest rate  $R$  be 1.02. As a result we have  $\sqrt{\beta^{-1} - R} \approx 0.05$ , which is negligibly small.<sup>21</sup> Defining  $\gamma \equiv \beta \frac{\sigma_E}{\sigma_d}$  and  $\varepsilon_{i,t} \equiv \epsilon_{i,t}^1 \epsilon_{i,t}^2$  to be the product of two independent random variables that follow a standard normal distribution and recalling that  $T_t = \frac{\tau_t W_t}{n}$ , the final law of motion can be further simplified to

$$w_{i,t} = (1 + \gamma \varepsilon_{i,t}) (1 - \tau) w_{i,t-1} + \tau_t,$$

which leaves  $\gamma$  as the only free parameter of our model.

---

<sup>21</sup> Note that a positive demand for the risky assets requires  $R < 1 + d$ . Stationarity of aggregate wealth furthermore demands for  $R < \beta^{-1} < 1 + d$  i.e.,  $\sqrt{\beta^{-1} - R} > 0$  but small. Furthermore, the variance of  $\epsilon_{i,t}^E$  is already relatively small. Rewriting  $\epsilon_{i,t}^d$  in terms of a standard normal reveals that the term is relatively small.



## 4.4 Analytical Results

This section aims to enrich our understanding of the process that generates the wealth distribution by finding a closed form solution for the stationary distribution as well as for the transition dynamics. In order to do so, we have to overcome some technical obstacles.

The portfolio returns, a product of two standard normal variables, follow a so-called *product-normal distribution*. To obtain a closed form solution, we have to transfer this distribution to another distribution that is easier to handle analytically.

**Proposition 2.** *The first three moments of the product normal distribution and the Laplace distribution with shape parameter of  $\lambda = \sqrt{0.5}$  are equal.*

*Proof.* See Appendix C.1.1. □

The Laplace distribution is very handy in our context for identifying a closed-form solution. The individual law of motion (LOM) has to be rewritten in continuous time in order to solve the Fokker-Planck equations which allows us to identify the cross-sectional distribution in terms of the free parameters  $\gamma$  and  $\tau$ . It would read as

$$dw_{i,t} = \tau(\bar{w} - w_{i,t})dt + (1 - \tau)\gamma w_{i,t}dNP, \quad (4.4.1)$$

for which  $NP$  is the noise following the product-normal distribution and  $\bar{w} = W/N$  the mean level of wealth. In order to retrieve a closed-form solution we transform this to the Laplace distribution using the scaling factor  $\lambda$  which we just introduced. The equation thus reads

$$dw_{i,t} = \frac{1}{\lambda}\tau(\bar{w} - w_{i,t})dt + \frac{1}{\lambda}(1 - \tau)\gamma w_{i,t}dL, \quad (4.4.2)$$

for which  $L$  signifies Laplace distributed noise.

**Proposition 3.** *Using Itô's lemma as a second-order approximation, disregarding redistribution for the very rich, and solving the Fokker-Planck equation, the stationary cross-sectional distribution*

can be expressed as a double Pareto distribution with an inverse shape parameter  $\alpha$ :

$$\alpha = 1 + \frac{\sqrt{2}\tau}{\gamma^2(1-\tau)^2}. \quad (4.4.3)$$

*Proof.* For the assumption of overall wealth  $W$  being identical to the number of agents  $N$ , the mean wealth is  $\bar{w} = \frac{W}{N} = 1$ . In fact, this process is a mean reversion process and the redistributive taxation makes individuals converge to the overall mean if noise is absent. The pace of convergence increases with the level of taxation  $\tau$ . Let us define the log of wealth  $\hat{w}_{i,t} = \log(w_{i,t})$  and apply Itô's lemma. Thus the equation reads

$$d\hat{w}_{i,t} = \left( -\frac{\tau}{\lambda} + \frac{\tau\bar{w}}{\hat{w}_{i,t}\lambda} - 0.5(1-\tau)^2\frac{\gamma^2}{\lambda^2} \right) dt + \frac{1}{\lambda}(1-\tau)\gamma dL.$$

Unfortunately, a complete closed-form solution is not feasible. We want to focus on the rich, whose (log) wealth  $\hat{w}_{i,t}$  is far higher than the mean wealth ( $\hat{w}_{i,t} \gg \bar{w}$ ). Thus, we can disregard the effect of redistribution. The equation simplifies to

$$d\hat{w}_{i,t} \approx \left( -\frac{\tau}{\lambda} - 0.5\frac{\gamma^2}{\lambda^2}(1-\tau)^2 \right) dt + \frac{1}{\lambda}(1-\tau)\gamma dL = -\mu dt + \delta dL,$$

with a diffusion term  $\delta \equiv \frac{1}{\lambda}(1-\tau)\gamma$  and a drift  $\mu \equiv \frac{\tau}{\lambda} + 0.5\frac{\gamma^2}{\lambda^2}(1-\tau)^2 = \frac{\tau}{\lambda} + 0.5\delta^2$ . As shown in Toda (2012), the Laplace distribution can be modeled by

$$dL = -\kappa \text{sign}(L)dt + \delta dB$$

with  $B$  being the standard Brownian motion and  $\text{sign}$  representing the sign function. Using this relationship we can rewrite the overall equation as

$$d\hat{w}_{i,t} = \begin{cases} \mu dt + \delta dB & \hat{w}_{i,t} < 0 \\ -\mu dt + \delta dB & \hat{w}_{i,t} > 0. \end{cases} \quad (4.4.4)$$

The cross-sectional distribution can be found by solving the so-called Fokker-Planck equation<sup>22</sup>

$$\frac{\partial f(\hat{w}, t)}{\partial t} = -\frac{\partial}{\partial \hat{w}} (\mu f(\hat{w}, t)) + 0.5 \frac{\partial^2}{\partial \hat{w}^2} (\delta^2 f(\hat{w}, t)).$$

We consider the stationary distribution ( $\frac{\partial f(\hat{w}, t)}{\partial t} \stackrel{!}{=} 0$ ). The solution is well-known (Karlin and Taylor, 1981, p. 221) and given by

$$f(\hat{w}) = 0.5\alpha \exp(-\alpha|\hat{w}|). \quad (4.4.5)$$

For our case we have

$$\alpha = \frac{2\mu}{\delta^2} = 1 + \frac{\sqrt{2}\tau}{\gamma^2(1-\tau)^2}. \quad (4.4.6)$$

As shown in Toda (2012), it is easy to transfer the Laplace distribution to a symmetric double Pareto distribution. In fact, if  $\hat{w}$  follows the described Laplace distribution, wealth  $w = \exp(\hat{w})$  is given by the probability density function

$$f(w) = \begin{cases} 0.5\alpha w^{-\alpha-1} & w \geq M \\ 0.5\alpha w^{\alpha-1} & 0 < w < M. \end{cases}$$

The mode of this double Pareto distribution is given by  $M = 1$ .<sup>23</sup> □

In fact, the complete distribution is characterized by the single value  $\alpha$ . Thus, other measures regarding inequality can be derived starting from this assumption.

**Proposition 4.** *The stationary ( $t \rightarrow \infty$ ) share  $s^x(\tau, \infty)$  of the top  $x$  (e.g. the top 1% implying  $x = 0.01$ ) wealth holders is given by*

$$s^x(\tau, \infty) = (1 + \alpha^{-1})0.5^{1/\alpha} x^{1-1/\alpha}, \quad (4.4.7)$$

---

<sup>22</sup> The latter is frequently referred to as Kolmogorov forward equation. The terms can, however, be used interchangeably.

<sup>23</sup> The cumulative probability distribution takes the following form:

$$F(w) = \begin{cases} 1 - 0.5w^{-\alpha} & w \geq 1 \\ 0.5w^\alpha & 0 < w < 1. \end{cases}$$

for which  $\alpha$ , as above, is implicitly a function of taxes  $\tau$  and  $\frac{\partial}{\partial \alpha} s^x(\tau, \infty) < 0$  for realistic case of  $\alpha > 1$ .

*Proof.* See Appendix C.1.1. □

The same rationale can also be used to derive a closed form for the Gini coefficient.<sup>24</sup> This implies that high  $\alpha$  is accompanied by low inequality. This very neat result has some strong implications. First of all, without taxation  $\tau = 0$  the tail-coefficient is  $\alpha = 1$ , identical to Zipf's law. In fact, the Gini coefficient then takes the value of  $Gini(w) = 1$  and  $s_x(\tau) = 1$  for all  $x \in (0, 1]$ , implying total inequality. Thus, in a laissez-faire economy without government intervention, there is no finite level of inequality. In general, inequality increases ( $\alpha$  decreases) with  $\gamma$  while decreasing with taxation  $\tau$ . For the extreme case of  $\tau \rightarrow 1$  - which can be thought of as a socialist society - we would have  $\alpha \rightarrow \infty$ , and thus have a Lorenz-curve identical to the 45-degree line and thus no inequality at all.

Note that our proof heavily relies on second-order approximations.<sup>25</sup> This is, however, not problematic for realistic values of  $\alpha < 2$ , for which only the first two moments exists.<sup>26</sup>

We can also make a statement about the convergence speed.

**Proposition 5.** *The convergence to the stationary distribution is given by*

$$\|f(w, t) - f(w, \infty)\| \sim \exp(-\phi t), \quad (4.4.9)$$

---

<sup>24</sup>The closed-form value for the Gini coefficient is given by

$$Gini(w) = \frac{3\alpha}{4\alpha^2 - 1}, \quad (4.4.8)$$

and decreasing with  $\alpha$  for realistic case of  $\alpha > 1$ . The proof can be found in Appendix C.1.1.

<sup>25</sup>In fact, the Fokker-Planck equation is also only a second-order approximation to the more general Master equation. Moreover, the use of Itô's lemma is only a second-order approximation. For noise generated by a Brownian motion (rather than the product-normal distribution) it would still hold exactly.

<sup>26</sup>As a result, we can only model degrees of inequality with  $s^{0.01} > s^{0.01}(a = 2) \approx 10.6\%$  using the closed-form solutions.

with an average convergence speed of

$$\phi = (0.5\gamma(1 - \tau)\alpha)^2. \quad (4.4.10)$$

*Proof.* The convergence speed is given by  $\phi = \frac{\mu^2}{2\delta^2}$ . See proposition 1 of Gabaix et al. (2016).  $\square$

This implies a half-life of  $t_{0.5} = \frac{\ln(2)}{\phi}$ . In fact the effective taxation  $\tau$  not only decreases steady-state inequality, but also increases the speed of convergence to the latter. This also means that there is an asymmetry in the convergence. The increase of inequality for low taxes is slower than the decrease after high tax rates. Thus, the positive message for the policy maker is that it is faster to come down to lower inequality rather than to increase the level of inequality.

**Proposition 6.** *The top-shares evolve according to an autoregressive process of first-order with*

$$s_t^x = \rho_t s_{t-1}^x + (1 - \rho_t) s^x(\tau_t, \infty), \quad (4.4.11)$$

and  $\rho_t = \exp(-\phi_t)$  for the average convergence speed  $\phi_t = \phi(\tau_t)$  as defined in Equation (4.4.10).

*Proof.* See Appendix C.1.1.  $\square$

Hence, the share owned by the fraction  $x$  of the population is a linear combination of last period's share and the share of the top  $x$  of the stationary distribution given the tax rate  $\tau_t$  at each time  $t$ .

## 4.5 An Empirical Application

In this section we feed empirical time series for taxes into our model's LOM in Equation (4.3) and attempt to match the data on wealth inequality. We start our investigation in the year 1900. From this year on (Piketty, 2014, chapter 14) provides data for the top income and top inheritance taxes.<sup>27</sup> Our approach is less rich than, for example Hubmer et al. (2016) or Kaymak and Poschke

---

<sup>27</sup>Note that wealth inequality data for France is already available from 1807. Yet, the data frequency is very low and moreover there is little variation in time.

(2016), yet allows for analytical tractability.<sup>28</sup> Rather than only focusing on the USA, we have a broader view by including the UK as well and - in order to contrast the Anglo-Saxon evolution - also consider France and Sweden.

We take the available tax series on capital gains taxes or, if these are unavailable, top marginal income taxes as a proxy for the wealth taxes introduced in Section 4.3. For the United States, tax data on capital gains taxes are available from the *Tax Policy Center*<sup>29</sup>, while for the other countries we have to rely on the marginal top income taxes by Piketty (2014). Moreover, we are only provided with the top capital income tax rate but not with the average tax rate.

**Proposition 7.** *Usually capital taxes only apply to the additional, positive returns on wealth and not on wealth itself or on losses. We approximate the gross-wealth tax  $\tau$  - given a capital gains tax  $\theta_r$  - by finding a  $\tau$  such that the expected value of after tax returns from a capital gains tax and after tax returns of a gross-wealth tax are equal. Gross-wealth taxes are then given by*

$$\tau = \frac{1}{2}\theta_r\gamma\lambda. \quad (4.5.1)$$

*Proof.* Dropping time subscripts for taxes, the after-tax returns given the capital income tax  $\theta_r$  are

$$\bar{R}_{\theta_r} = 1 + \begin{cases} (1 - \theta_r)\gamma\varepsilon_{i,t} & \text{if } \varepsilon_{i,t} > 0 \\ \gamma\varepsilon_{i,t} & \text{if } \varepsilon_{i,t} \leq 0. \end{cases}$$

To use the LOM in Equation (4.4.11) we approximate  $\tau$  given  $\theta_r$  by finding a  $\tau$  such that the expected value of  $\bar{R}_\tau$  equals the expected value of  $\bar{R}_{\theta_r}$ . Then, given that  $\varepsilon_{i,t}$  approximately follows a Laplace distribution with scale  $\lambda = \sqrt{0.5}$ , the expected value  $E[\varepsilon_{i,t}|\varepsilon_{i,t} \leq 0]$  is the mean of an

---

<sup>28</sup> Note that the approach of Aoki and Nirei (2017) is also analytically tractable to a high extent.

<sup>29</sup> The data is available at <http://www.taxpolicycenter.org/statistics/historical-capital-gains-and-taxes>.

exponential distribution with inverse scale  $\lambda$ , which is again  $\lambda$ . Then

$$E\bar{R}_\tau = E\bar{R}_{\theta_r}$$

$$E\{(1-\tau)(1+\gamma\varepsilon_{i,t})\} = 1 + \gamma P(\varepsilon_{i,t} \leq 0)E[\varepsilon_{i,t}|\varepsilon_{i,t} \leq 0] + (1-\theta_r)\gamma P(\varepsilon_{i,t} > 0)E[\varepsilon_{i,t}|\varepsilon_{i,t} > 0]$$

$$1 - \tau = 1 - 0.5\gamma\lambda + 0.5(1 - \theta_r)\gamma\lambda$$

$$\tau = \frac{1}{2}\theta_r\gamma\lambda,$$

where  $P(\varepsilon_{i,t} > 0)$  denotes the probability that  $\varepsilon_{i,t}$  is positive.<sup>30</sup> □

Unfortunately, time series on capital or wealth taxation are unavailable for the UK, France and Sweden, which is why we focus on the data on income taxation as a proxy. In order to account for income instead of capital taxation, we cannot directly use the available data as  $\theta_{r,t}$ , as we do in the case for the US simulations. Using the top marginal income taxes directly will potentially imply too much taxation for various reasons. Frequently, income taxes are far more complicated than the capital gains taxes considered here. Income taxes are progressive to a varying degrees while the above approximation applies the top marginal tax rate to all agents independent of the magnitude of their capital gains or income level. Also, capital gains taxes are, for most countries, considerably lower than income taxes due to higher mobility of capital (as compared to labor) giving countries an incentive to offer lower capital gain taxes in the international tax competition. Furthermore, both income and capital gains taxes provide different forms of exemptions that can not be modeled in detail if we want to make use of our closed form solution. Lastly, we do not adjust for the possibility of tax evasion and tax avoidance.<sup>31</sup>

For this reason we scale down the available series of top income taxes by multiplying it with a scale parameter  $0 < \eta \leq 1$  to adjust for progressivity, exemption levels, the difference between

---

<sup>30</sup> In fact, this also approximates the conversion of annual tax rates to a quarterly calibration.

<sup>31</sup> Analyzing evidence from leaked tax evasion documents, Alstadsaeter et al. (2017) show an increase of tax evasion with wealth rising up to 30% for the top 0.01% of the wealth distribution.

capital gains taxes and income taxes, and tax avoidance. We define

$$\tau_t = 0.5\lambda\gamma\eta\theta_{r,t}. \quad (4.5.2)$$

Our second free parameter is  $\gamma = \beta \frac{\sigma_E}{\sigma_d}$ . The finance literature usually uses a value of  $\sigma_d$  in the range from 0.08 to 0.3 (Campbell and Viceira, 2002). Estimates for return disagreement are hard to gather. Greenwood and Shleifer (2014b) present values ranging between 1% and 3%. Combining these values implies an estimate for  $\gamma$  in the range from 0.1 to 0.375.<sup>32</sup>

To obtain some intuition, let us plug in Equation (4.5.2) into the closed-form solution from Equation (4.4.11), and approximate  $(1 - \tau) \approx 1$ . Then

$$s_t^x = e^{-\phi_t} s_{t-1}^x + (1 - e^{-\phi_t})(1 + 1/\alpha_t)0.5^{1/\alpha_t} x^{1-1/\alpha_t},$$

with

$$\begin{aligned} \alpha(\theta_{r,t}) &\approx 1 + \frac{0.5\theta_{r,t}}{\gamma}, \\ \phi(\theta_{r,t}) &\approx \frac{1}{4}\gamma^2 + \frac{1}{4}\theta_{r,t}\gamma + \frac{1}{16}\theta_{r,t}^2, \end{aligned}$$

for which follows that *if the system is at the steady state*

$$\frac{\partial s_t^x}{\partial \gamma} > 0 > \frac{\partial s_t^x}{\partial \theta_{r,t}} \quad \text{and} \quad \frac{\partial |\Delta s_t^x|}{\partial \theta_{r,t}}, \frac{\partial |\Delta s_t^x|}{\partial \gamma} > 0.$$

The weight on the most recent value  $s_{t-1}^x$  decreases in the transition speed, which depends positively on taxes  $\theta_{r,t}$  and dispersion  $\gamma$ . This means that in terms of inequality dynamics, an increase in the standard deviation of portfolio returns is a complement to an increase in taxes and will speed up dynamics. However, in terms of inequality levels these two have opposing effects: an increase in taxes  $\theta_{r,t}$  decreases the stationary level of inequality, while a higher value of  $\gamma$  will increase it.

---

<sup>32</sup> Keep in mind that for our quarterly calibration we set  $\beta = 0.99 \approx 1$ .



We use this LOM (without the approximation) to estimate  $\gamma$  (and  $\eta$  if necessary) for each country by using numerical optimization routines to minimize the squared difference between a time series created by the LOM in Equation (4.4.11) and the available data. Denote as  $s_t^{0.01}(\{\tau\}_0^t, S)$  the 1% share at  $t$  of a series initialized with  $S \in (0, 1)$  and fed with the tax series  $\{\tau\}_0^t$ . This model-generated time series is compared to the empirical top 1% share at time  $t$ , denoted as  $\hat{s}_t^{0.01}$ . Then

$$\{\gamma, \eta\} = \arg \min \sum_{t=0}^T \left[ s_t^{1\%}(\{\tau\}_0^t, \hat{s}_0^{1\%}) - \hat{s}_t^{1\%} \right]^2. \quad (4.5.3)$$

The resulting estimates are summarized in Table 4.1 and the (analytic) trajectories can be found in Figure 4.2.<sup>33</sup> We will discuss these results in detail below.

	US	UK	SE	FR
$\gamma$	0.3523	0.0398*	0.2012	0.1341
$\eta$	–	0.2484	0.3709	0.2372

Table 4.1: Parameter estimates used for closed-form low-of-motion and simulations. For the UK the analytic solution underestimates  $\gamma$ , which is why for the simulations  $\gamma = 0.082$  is used.

As reference, we also show simulations using Equations (4.3) and (4.5.2), and given the empirical tax rate at each time  $t$ . These simulations are important in so far, that they allow us to identify the 95% confidence interval of our model, which is useful for evaluating its empirical fitness. Since the real world is a stochastic environment with a finite number of agents but the closed form assumes an infinite number of agents, we can assess the degree of discrepancy.

For simulations, we created an initial distribution Pareto distribution that satisfies  $W = \sum w_0 = N = 50,000$  and matches the top 1 and top 0.1 percentiles from the WID database for the respective moment in time.<sup>34</sup> Starting from this distribution, the series follows the model outlined above

<sup>33</sup> For further comparison, we also quantified the fitness using mean absolute percentage errors (MAPE) and the Pearson correlation coefficient (PCC):

	US	UK	FR	SE
MAPE	0.0639	0.0782	0.0904	0.0858
PCC	0.8777	0.9948	0.9579	0.9827

<sup>34</sup> This can be done easily by using the well-known relationship with the share of the top  $x\%$  given by  $s^x = x^{1-1/a}$ . We can then estimate the Pareto coefficient to fit the top 1% and the top 0.1%, create a distribution and check the fit with the empirical values. We do not intend to match the lower 50% interval since our model is one of investment and frictionless financial markets. These assumptions do not hold for the bottom of the wealth distribution.

while feeding in the tax series as explained above. Robustness checks for the variation of the free parameter  $\gamma$  are presented in Appendix C.2.

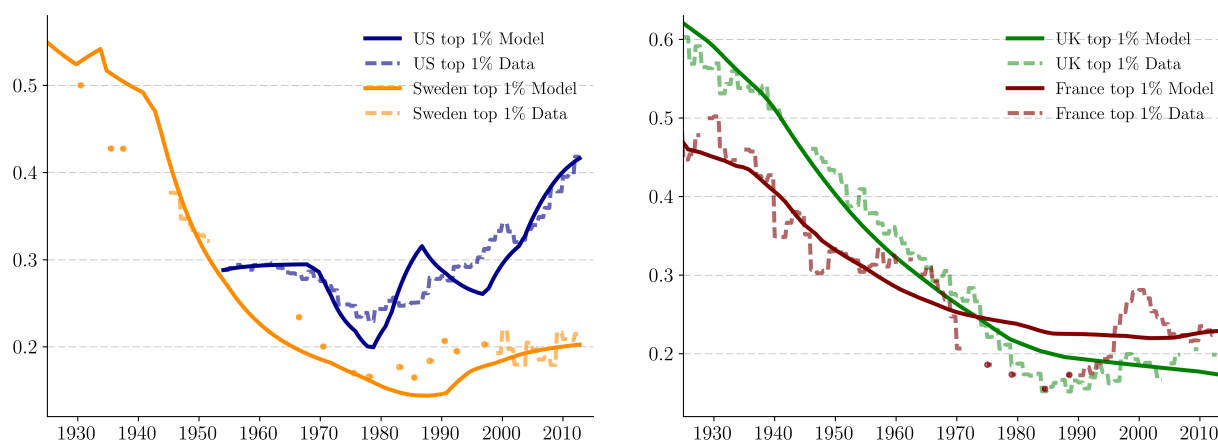


Figure 4.2: Trajectories using the law-of-motion in closed form (Equation 4.4.11) and the parameters from Table 4.1. Each time series is initialized with the top-share from the empirical data and then uses the available series of (capital) income taxes as the single input.

### 4.5.1 The United States

In the United States, individuals generally pay income tax on the net of their capital gains. There are, however, a reasonable number of exemptions, depending on investment duration, net-worth and status. The series in use here represents the maximum tax rate on long-term gains obtained from the US Tax Foundation. This has the great advantage that it is a tax explicitly and only on capital gains, which spares the need to use (and estimate)  $\eta$  to scale down top marginal income taxes. Yet, we note that reducing a complicated system of tax progression to just one number bears the risk of misalignment.

In Figure 4.3, we plot our closed form solution for the 1% top shares and the US calibration. The dashed line stands for the analytic result of the final cross-sectional distribution after convergence and given the tax rate  $\theta_{r,t}$  at time  $t$ . The solid line is the analytic LOM. For this reason the final analytic top-shares *jump* with each change in the tax rate and the numerical simulation follows more slowly, in line with the data.

The results, together with the series of taxes, are shown in Figure 4.4. The dashed blue/orange line depicts the original data while we add the median as well as the 95% confidence intervals from our simulations. Median and intervals from 200 simulations are calculated starting from the same distribution and using different random seeds. We initialize our simulation for the US in 1954 because from this year on capital gains taxes are available on an annual basis. In Appendix C.4 we presents histograms of the overall simulated data and compare it to the closed-form solutions.

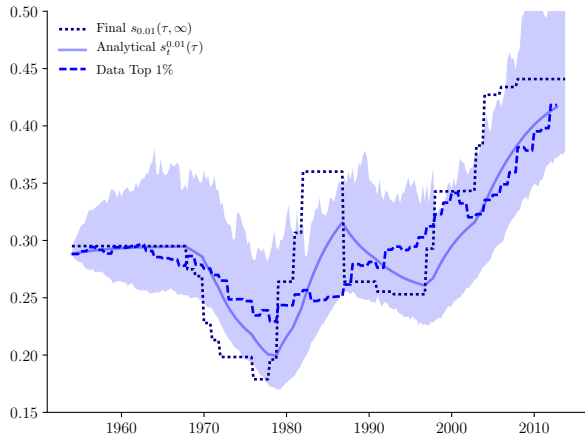


Figure 4.3: Closed form LOM  $s_t^{0.01}(\tau_t)$  (solid line) and closed form stationary share  $s^{0.01}(\tau_t, \infty)$  (dotted line) taking tax series and the calibration for the USA together with the empirical evidence (dashed line) and the 95% interval of the respective simulations.

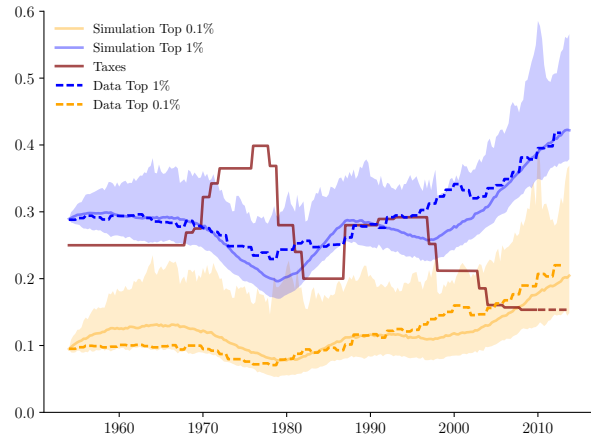


Figure 4.4: Simulations starting from the wealth distribution in 1954 for the USA, using only the respective time series of US capital gain taxes as input ( $\gamma = 0.3523$ ).

The value of  $\gamma$  is 0.3523, which implies that the standard deviation of forecasting errors is yet very small compared to the real standard deviations of returns and hence also lies in a sensible range. Thus, our value of  $\gamma = \beta \frac{\sigma_E}{\sigma_D}$  for the USA would be in line with relatively high disagreement variance and low return variance. Our simulation matches the data very well with the time series of the data, always lying within the 95% interval of our simulation and the median sticking close to the real time series.

In the late 1950s the series seems to be stable around the steady state given constant taxes. After successively increasing capital gain taxes starting in the late 1960s, inequality decreases up until the Reagan period, in which taxes return to the previous level. In particular, the levels as well as the

responses to this tax increase are well matched while simultaneously providing realistic transition dynamics. The relaxation of taxes is accompanied by an immediate increase in wealth inequality both in the data and predicted by our model. While, after a short dip, taxes return to their level from the 1950s in the late 1980s, inequality stabilizes at the same postwar-level in the data as well as in the model. Finally, the tax decreases in the late 1990s and the early 2000s initiate convergence to a steady state that is yet unreached. It is furthermore noteworthy that, even though the 95% intervals indicate a considerable variation among simulations, the amount of total variation remains almost constant and the movement of the outer boundaries coincides with the movements in the data.

## 4.5.2 The United Kingdom

Let us now turn to the United Kingdom. The tax system in place was introduced in 1965. However, we start our investigation back in 1913. Thus, we rely on the only available tax series dating back that far, which is the series on top marginal income taxes provided by Piketty (2014). In fact, the actual capital gains tax rate in 2010 (depending on income) ranges from 18% up to 28% while allowing for several exemptions. This, again, leaves a considerable degree of freedom when attempting to summarize these tax regimes by one number per period, supposed to address only the top shares.

Figures 4.5 and 4.6 show our model results and the simulation for the UK starting in 1904 and document the income taxation for each period, taken from Piketty (2014). Compared to the US we opt for a  $\gamma$  of 0.082.<sup>35</sup> The ratio fits in line with the reported evidence for expectation disagreement (Greenwood and Shleifer, 2014b) and return variance (Campbell and Viceira, 2002). Compared to the USA the value of  $\gamma$  is lower, implying less disagreement among traders. In fact, in the sample

---

<sup>35</sup> For the US and France we used the same values for our simulation that were identified by using the closed form solution. This is not feasible for the UK, since the second-order approximations used for our analytic results becomes imprecise when  $\gamma$  is very small. In particular, for  $\alpha > 2$  the closed-form results fail to hold. Moreover, our closed-form solution slightly underestimates  $\gamma$  because it understates inequality when taxes are relatively high. For this reason, for the UK and Sweden, we repeated the estimation procedure using the median of the simulation instead of the analytic LOM.

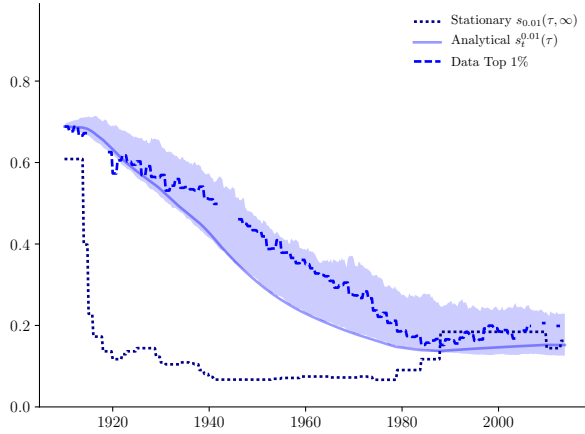


Figure 4.5: Closed form LOM  $s_t^{0.01}(\tau_t)$  (solid line) and closed form stationary share  $s^{0.01}(\tau_t, \infty)$  (dotted line) taking tax series and the calibration for the UK together with the empirical evidence (dashed line) and the 95% interval of the respective simulations.

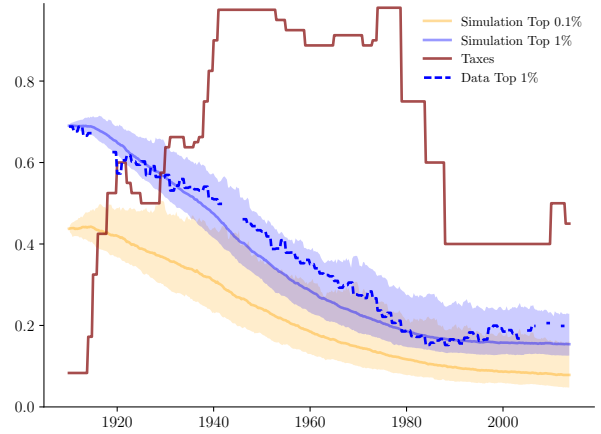


Figure 4.6: Simulations starting from the wealth distribution in 1913 for the UK, using only the respective time series of UK top income taxes as input.  $\gamma = 0.082$  and  $\eta$  as in Table 4.1.

of the countries considered in this study the United Kingdom is the one with the lowest wealth inequality.

In accordance with the data, the model predicts an overall decrease in wealth inequality until the Thatcher years, in which - similar to the USA - taxes decreased considerably. This is reflected by a gradual increase and rise in inequality, both in the data and predicted in the model. Again, not only the levels but also the slow transitions are well matched. As shown in Figure 4.5 the stationary level of inequality in the periods of the tax hike until the 1980s is substantially below the transition level owing to low transition dynamics. This is especially surprising because one would expect to obtain a worse fit given that marginal top income taxes are not an optimal proxy for capital income taxes. Yet, after initiation of the distribution given only one relatively narrow bounded parameter, our model provides a good fit for over 100 years of data.

### 4.5.3 Sweden

Let us have a look at Sweden. The usual calibration procedure produces a value of  $\eta = 0.3709$  higher than in the other considered cases. The estimate for  $\gamma$  is 0.2186, lower than the value for

the USA, but higher than the value for the UK. The results are presented in Figures 4.7 and 4.8. Similar to the Anglo-Saxon economies, there was an increase in taxes and thus a decrease in wealth inequality until the 1980s in which the trend reversed. Roine and Waldenström (2015) discuss the evolution of wealth inequality in Sweden in detail. The strong increase in taxation after the Second World War was a reaction of the ruling Social Democratic Party to crowd out the nascent Communist movement in Sweden. Sweden had a wealth tax in place from 1911 until 2007, when it was abolished when a center-right government came to take over the government responsibilities. Due to this fact a pure wealth tax is not available for the complete sample. Thus - and therefore in line with our approach for the UK and France - we approximate the wealth tax by the income tax.

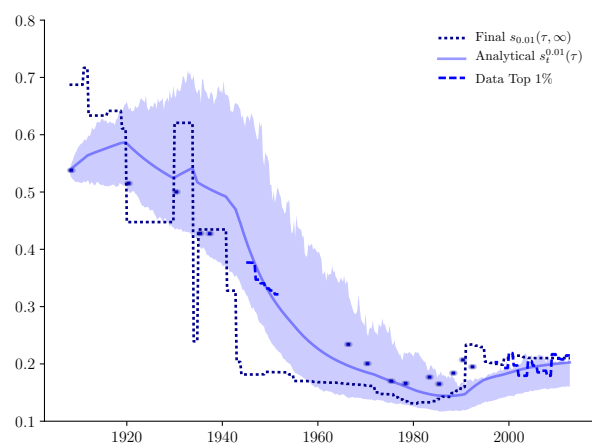


Figure 4.7: Closed form LOM  $s_t^{0.01}(\tau_t)$  (solid line) and closed form stationary share  $s^{0.01}(\tau_t, \infty)$  (dotted line) taking tax series and the calibration for Sweden together with the empirical evidence (dashed line) and the 95% interval of the respective simulations.

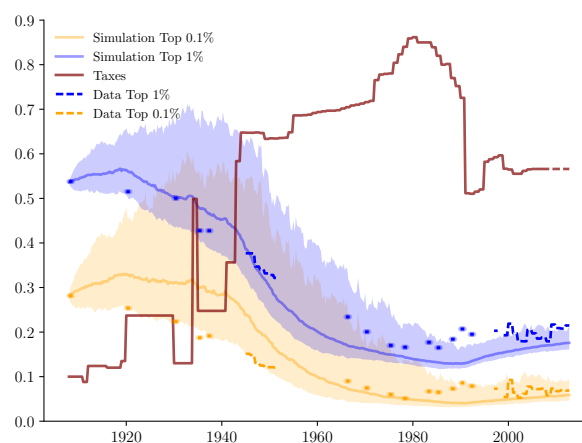


Figure 4.8: Simulations starting from the wealth distribution in 1908 for Sweden, using only the respective time series of Swedish top income taxes as input. Parameters as in Table 4.1.

As documented in Du Rietz and Henrekson (2014) the top marginal wealth tax rate increased until the 1980s, decreasing afterwards in line with the evolution of top income tax rates. In general, Du Rietz and Henrekson (2014) presents a detailed description about the evolution of taxation in Sweden including capital gain and wealth taxes. Yet, for reasons of consistency with the data on wealth inequality we use the tax data provided by Lundberg and Waldenström (2017).

Once again we are able to match the decrease in wealth inequality until the 1980s (cf. figure 4.8). After a substantial tax decrease top wealth inequality is on the rise again. As furthermore presented in the closed form solution depicted in figure 4.7 it is, however, already close to its stationary value. As before, the transition dynamics are well in line with empirical evidence.

#### 4.5.4 France

We also want to compare the two Anglo-Saxon economies and Sweden that underwent deep regulatory changes in the 1980s with France as an example for a central European economy. In France, capital gains taxes, in addition to taxes on sales of financial instruments, are generally subject to the marginal tax rate plus some social contributions. Certain deductions might be feasible depending again on holding duration and other individual indicators. Thus, we again keep in mind that these deductions, together with potential changes in tax progression, can not be captured well by the one-number-summary of each periods tax regime.

We again start our investigation in the year 1917. From this year on data for the top income and top inheritance taxes are available (Piketty, 2014, chapter 14). Similar to the UK, the empirical time series on capital gain taxes is not available. The fitted value of  $\gamma$  amounts to 0.1341, slightly lower than for the case of Sweden, but higher than the value for the UK, while  $\eta$  is estimated to be 0.2372. Basically, we see an increase in taxes until World War II. Top income taxes leveled out afterwards only to slightly decrease after the 1980s. Yet, the decrease in taxes is not as pronounced as in the other countries covered, which explains the lower degree of wealth inequality.

Figure 4.10 presents simulation results. While in the beginning our model is able to match the data fairly well, the fit from the 1970s until the 1990s is rather loose. However, in this period only a few data points are provided and most of the series still remains within our 95% interval. It is furthermore noteworthy that the empirical series shows relatively large fluctuations, which is also reflected in relatively abrupt changes in the tax rate. The main driving force of this result is that the distribution starts off from a steady state with very low taxation, and then after the introduction of relatively high tax rates converges slowly to the lower steady state.

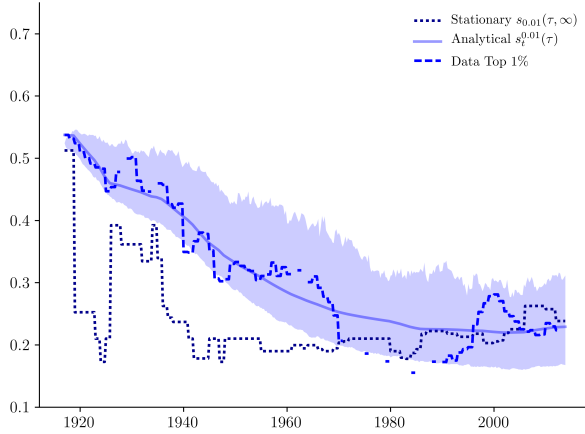


Figure 4.9: Closed form LOM  $s_t^{0.01}(\tau_t)$  (solid line) and closed form stationary share  $s^{0.01}(\tau_t, \infty)$  (dotted line) taking tax series and the calibration for France together with the empirical evidence (dashed line) and the 95% interval of the respective simulations.

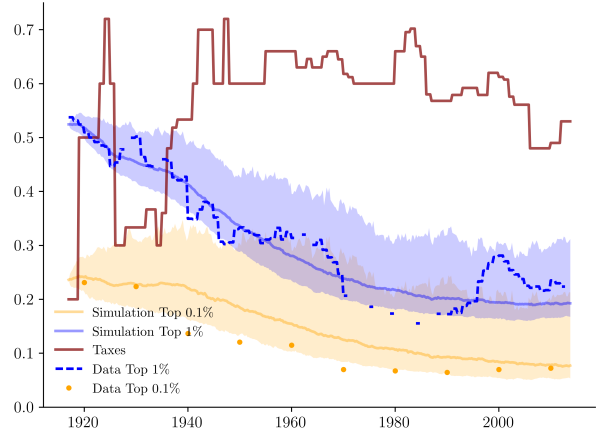


Figure 4.10: Simulations starting from the wealth distribution in 1917 for France, using only the respective time series of French top income taxes as input. Parameters as in Table 4.1.

## 4.6 Validation and Projections 2040

We validate our model through out-of-sample testing. Instead of estimating the free parameters  $\gamma$  (and potentially  $\eta$ ) for each country given the full available dataset, we only use a shorter sample from the history to determine  $\gamma$  and then initiate our model into an episode that does not lie within the sample. For this purpose we make use of our closed-form solution from Equation (4.4.11) and feed in only the available tax series. The model simulations can then be compared with the empirical data for the periods after the sample phase.

The quality of our out-of-sample predictions depends crucially on the properties of the sample period. For the US, our sample starts in 1956 as above. The years thereafter stand out as a period of very stable inequality. As we have argued above,  $\gamma$  has two effects on inequality dynamics. First, a lower (higher)  $\gamma$  is associated with a lower (higher) steady state level of inequality. Second, a lower (higher)  $\gamma$  is linked with slower (faster) transitions. Hence, if the sample contains a stationary period, the respective value of  $\gamma$  can be easily identified through the first criterion. The sensitivity and robustness with respect to this parameter is further discussed in Appendix C.2.



For the case of the USA, proper identification is even easier because the stationary period is - after an increase in taxes in the late 1960s - followed by a significant decrease in inequality. Furthermore, there is no need to estimate  $\eta$  since capital taxes are readily provided. For this reason our US dataset allows a good approximation with  $\gamma = 0.3497$  when using only a sample of 10 years (1956 - 1966). Increasing the sample size does not then have a significant effect on the estimate of  $\gamma$ . In Figure 4.11 (left) we show the fit of this simulation for different samples for the USA. For each sample, the fit is very close to our benchmark results from the previous section since the associated  $\gamma$  is very close to its value for the whole available dataset.

Unfortunately, such periods of stable inequality are absent in the data for the other countries. Since for the UK, France, and Sweden wealth inequality is constantly decreasing until the 1980s, given any subsample of this data it is only possible to identify that inequality is at a transient path, but this is insufficient to narrow down our parameters. This can be readily understood by the same argument as before. If inequality is high and taxes are high too, inequality will decrease until the stationary level is reached where convergence speed and final level both depend on  $\gamma$ . If the value of  $\gamma$  is low, the final level will be low, but convergence speed will be low as well. If the sample only includes a decreasing path of inequality, it is not possible to identify whether the increase is due to slow convergence to a low level, or fast convergence to a high level, as long as the final level is not part of the sample data. Further, for the mentioned three countries it is also necessary to identify  $\eta$ , which reflects a further complication since even if a period of stable inequality would be given, the estimation would be singular. The parameter  $\eta$  has a contrasting effect compared to  $\gamma$ . An increase (decrease) also leads to an increase (decrease) in transition speed, but to a lower (higher) level of inequality.

We focus on the parameter estimation for Sweden here and redirect estimation and discussion for the UK and France to the Appendix. In short, while our procedure shows mild success for the UK given the limitations discussed above, the erratic nature of the data for France does not allow a precise and liable estimation of the parameters. We also adjust our estimation procedure slightly, by iteratively estimating only one parameter at a time while using the latest estimate of

the others. This procedure is repeated until changes in the parameter estimates become marginal. This is to avoid the singularity problem of estimating both parameters simultaneously given a very small sample.<sup>36</sup>

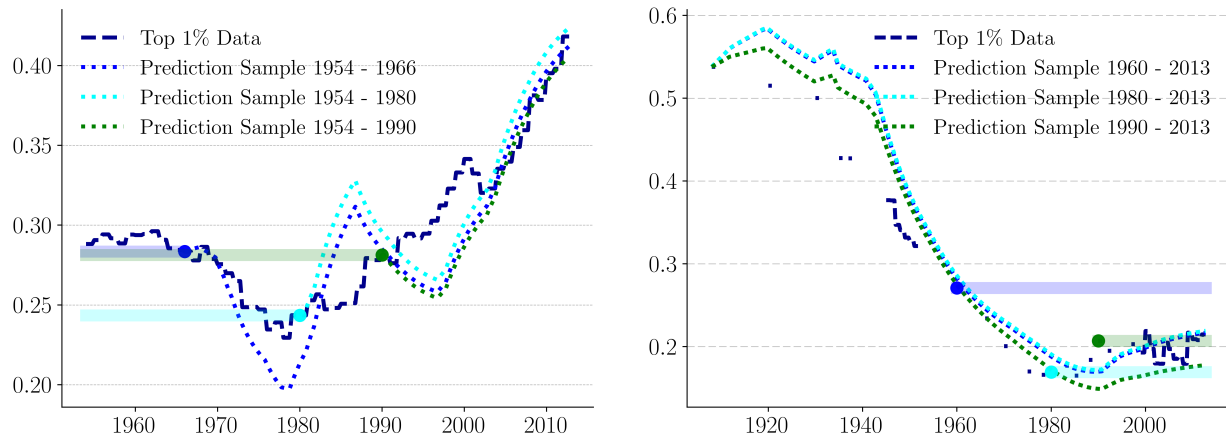


Figure 4.11: Out-of-sample tests for the US (left) and Sweden (right). The fat bars indicate the sample period.

Keeping in mind our discussion above, Sweden does not provide a stationary period either. To find a value for  $\gamma$  and  $\eta$  that captures the dynamics as precisely as in the previous section, the period that follows 1990 is essential. If we, however, for the reasons outlined above, only use this period and initialize the comparing simulation at  $t = 0$ , we are able to estimate  $\gamma$  and  $\eta$  such that they provide a fairly good fit for the out-of-sample period. In Figure 4.11 (center) we show out-of-sample predictions for Sweden in different sample sets ending in 2013.

As discussed the out-of-sample predictions work well for the USA implying that we can also make reasonable future forecasts for a long and very long run horizon. The same holds true for the last 20 years of Swedish data featuring a rather constant level of inequality helping us to identify the parameters. Using the calibration from the previous section and starting from a Pareto distribution, we first simulate the last 20 years for which we know the empirical shares and the time series of taxes. This then also serves as a second test for our calibration, which was targeted to match a longer horizon.

<sup>36</sup>In fact, this procedure then explicitly searches for local minima only, a compromise that we are willing to make in this context.

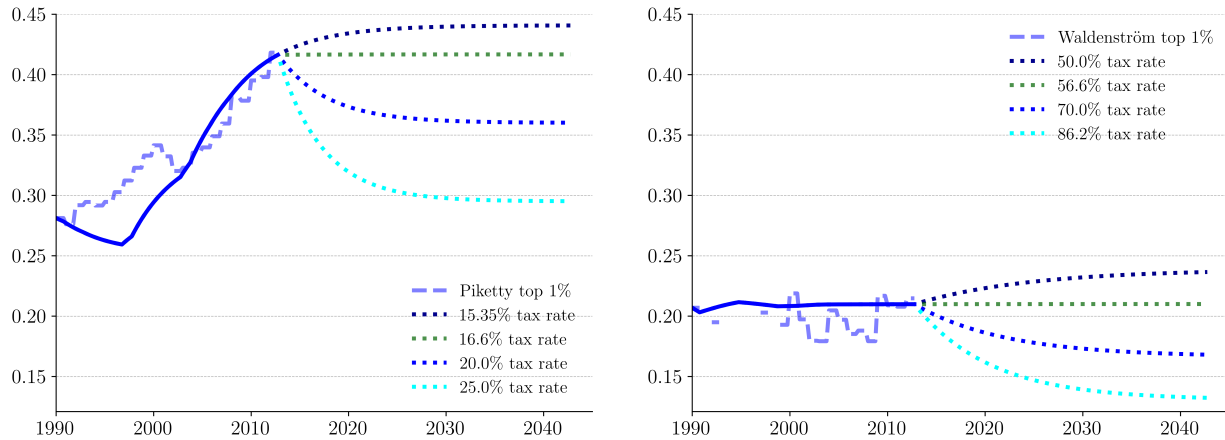


Figure 4.12: At the crossroads: projections for different tax regimes. For the US (left) the top dashed line projects the current tax rate, the second line shows that a rate of at least 16.6% would be necessary to keep inequality at the present level. A rate of 20% would partly reverse the trend. The bottom line shows a 25% rate. For Sweden a rate of 56.6% is required to keep inequality at the present level.

In Figure 4.12 (left) we show analytic forecasts given the US estimation and different tax regimes. A considerable increase in the capital gains tax rate to 25%, as for example currently prevailing in Germany would be necessary to reverse the trend and bring inequality back to the level of the 1990s. For the other countries, simulations and closed form solutions imply that, given timely tax rates, the stationary level of inequality is almost reached. Figure 4.12 (right) displays the (exemplary) behavior of Sweden, showing that for a tax rate of 56.6% wealth inequality is at a steady state level. Lower, respectively, higher tax rates will increase (decrease) the level of wealth inequality.

## 4.7 Conclusion

The main purpose of this work is to develop a simple, yet micro-founded model to explain the dynamics of wealth inequality given empirical tax series. Although a quite straightforward approach, this stands in contrast to the majority of the theoretical literature on wealth inequality which takes income inequality as a starting point.

We apply this method to the USA, the UK, France, and Sweden. For the USA we use capital taxes, while for the UK, Sweden and France marginal top income taxes are taken as a proxy. Due to the parsimonious nature of our model the degree of freedom to match the data is very limited. Nevertheless, our model matches the data surprisingly well, both in levels and also in transition speed. Our analytic results emphasize that the level and the transition speed of wealth inequality depend crucially on the degree of capital taxation, which is well in line with our numerical results quantitatively and qualitatively. We conclude that the given tax series have a very high explanatory power regarding the dynamics of wealth distribution over the last 70-100 years.

This also implies that the answer to the question of how to approach – and potentially reverse – the recent increase in wealth inequality that can be observed in developed economies is considerably simple. An increase in capital gains taxes, or alternatively a gross tax on wealth, as suggested in Piketty (2014), will very likely mitigate the issue and has the potential to upturn these trends. Our projections predict that, in particular for the USA - continuing on the present path of capital taxation - the gap between the rich and poor is expected to further increase.

There are two implications for future research. Although our model fits the data quite well, there are periods where it falls short of accounting for the data. First, we consider it important to identify whether the loose fit is due to bad tax data or to reasons that are exogenous to our model. Second, if these reasons are exogenous it is crucial to investigate them further.

Of course, the quality of the model's result severely hinges on the quality of the data. In particular, a better availability of data on wealth dispersion at higher frequencies would give better means to testing and improving our model and improve the understanding of the issue of wealth inequality in the 21st century.

# Appendix C

## C.1 Proofs of the Propositions

### C.1.1 Proof of Proposition 2

The product-normal distribution is treated extensively in Craig (1936). The probability distribution function is given by

$$f(z_{PN}) = \frac{1}{\pi} K_0(|z_{PN}|), \quad (\text{C.1.1})$$

with  $z_{PN} \equiv \epsilon^1 \epsilon^2$  with  $\epsilon^i \sim N(0, 1)$  and  $K_0$  being the modified Bessel-function of the second kind. The function is symmetric around the mean of zero and exhibits leptokurtic behavior. It is more appealing to write this using the Moment-Generating Function (MGF), which in this case is given by

$$M_{Z_{PM}}(t) = \frac{1}{\sqrt{1-t^2}}. \quad (\text{C.1.2})$$

Using this it is easy to show that the mean and skewness are zero, while the standard deviation is given by

$$SD(z_{PN}) = 1. \quad (\text{C.1.3})$$

This distribution is highly comparable to the Laplace distribution. For a zero-mean the probability density function of the latter is given by

$$f(z_L) = \frac{1}{2\lambda} \exp\left(-\frac{|z_L|}{\lambda}\right) \quad (\text{C.1.4})$$

for shape parameter  $\lambda > 0$ , having both a mean and a skewness of zero. The standard deviation of Laplace is

$$SD(z_L) = \sqrt{2}\lambda. \quad (\text{C.1.5})$$

The Laplace distribution is also very appealing as each half takes the form of an exponential function. The moment generating function of the Laplace distribution is

$$M_{Z_L}(t) = \frac{1}{1 - \lambda^2 t^2}. \quad (\text{C.1.6})$$

Comparing this with the MGF of the product-normal distribution it becomes obvious that the two are not identical. In fact, the sum of two product-normal variables follows a Laplace distribution.<sup>1</sup>

As a reasonable approximation we replace the product-normal with the Laplace distribution. To obtain the shape parameter  $\lambda$  that best approximates the standard normal product distribution we equalize the second order Taylor expansions of both MGFs around  $t = 0$ , which in fact is equivalent to choosing  $\lambda$  to match the first two moments of the function. This yields

$$\begin{aligned} \sum_{n=0}^2 \frac{\partial^n M_{Z_{PM}}(0)}{n! \partial t^n} (t-0)^n &= \sum_{n=0}^2 \frac{\partial^n M_{Z_L}(0)}{n! \partial t^n} (t-0)^n \\ 1 + \frac{t}{2} &= 1 + t^2 \lambda^2 \\ \lambda &= \sqrt{\frac{1}{2}} \approx 0.707. \quad \square \end{aligned}$$

---

<sup>1</sup>Using the MGF it is easy to show that if there are four independently distributed normal shocks with zero mean  $X_i \sim N(0, \sigma_i)$  and we have  $\sigma_1 \sigma_2 = \sigma_3 \sigma_4$  then  $X_1 X_2 + X_3 X_4$  follows a Laplace distribution with zero mean and  $\lambda = 1$ .

### Proof of Proposition 4

From the CDF given by

$$F(w) = \begin{cases} 1 - 0.5w^{-\alpha} & w \geq 1 \\ 0.5w^\alpha & 0 < w < 1, \end{cases} \quad (\text{C.1.7})$$

we can compute the Lorenz-curve with

$$L(F) = \frac{\int_0^F w(F')dF'}{\int_0^1 w(F')dF'}. \quad (\text{C.1.8})$$

This requires the inverse of the CDF which takes the following form

$$w(F) = \begin{cases} 2^{-\frac{1}{\alpha}}(1 - F)^{-\frac{1}{\alpha}} & F \geq 0.5 \\ 2^{\frac{1}{\alpha}}F^{\frac{1}{\alpha}} & 0 < F < 0.5. \end{cases} \quad (\text{C.1.9})$$

With integration by parts we can write

$$L(F) = \frac{\int_0^{0.5} w(F')dF' + \int_{0.5}^F w(F')dF'}{\int_0^{0.5} w(F')dF' + \int_{0.5}^F w(F')dF'}, \quad (\text{C.1.10})$$

for the relevant part  $F > 0.5$  (the *rich*). Integrating in the boundaries leads to

$$L(F) = \frac{\frac{\alpha}{\alpha+1}0.5 - \frac{\alpha}{1-\alpha}0.5 + 2^{-\frac{1}{\alpha}}\frac{\alpha}{1-\alpha}(1 - F)^{1-\frac{1}{\alpha}}}{\frac{\alpha}{\alpha+1}0.5 - \frac{\alpha}{1-\alpha}0.5}. \quad (\text{C.1.11})$$

Let us define the denominator  $z \equiv \frac{\alpha}{\alpha+1}0.5 - \frac{\alpha}{1-\alpha}0.5 = -\frac{\alpha^2}{1-\alpha^2}$ . Thus, we have

$$L(F) = 1 + \frac{2^{-\frac{1}{\alpha}}\frac{\alpha}{1-\alpha}(1 - F)^{1-\frac{1}{\alpha}}}{z} \quad (\text{C.1.12})$$

Following the logic of the Lorenz-curve the share  $s^x$  of the top  $x$  is given by

$$s^x = 1 - L(1 - x). \quad (\text{C.1.13})$$

Using the definition we can write

$$s^x = -\frac{2^{-\frac{1}{\alpha}} \frac{\alpha}{1-\alpha} (1-F)^{1-\frac{1}{\alpha}}}{z} = -\frac{2^{-\frac{1}{\alpha}} \frac{\alpha}{1-\alpha} (x)^{1-\frac{1}{\alpha}}}{z} = \frac{1+\alpha}{\alpha} 2^{-\frac{1}{\alpha}} x^{1-\frac{1}{\alpha}}. \quad \square \quad (\text{C.1.14})$$

It is important to point out the relationship to the standard Pareto-distribution, for which the share is given by  $s^x = x^{1-\frac{1}{\alpha}}$ . For the symmetric double Pareto it amounts to  $s = Kx^{1-\frac{1}{\alpha}}$  with  $K = \frac{1+\alpha}{\alpha} 2^{-\frac{1}{\alpha}} \geq 1$ . Note that for the special case of  $\alpha = 1$  respectively the inverse case  $\alpha \rightarrow \infty$  we have  $K = 1$ . In fact,  $K$  is a u-shaped function with a local maximum at  $\alpha = \frac{\ln(2)}{1-\ln(2)} \approx 2.26$  amounting to  $K = \frac{2}{e \ln(2)} \approx 1.06$ . Thus, the top share is slightly higher than in the standard Pareto-case. Yet, the standard Pareto-approach fares quite well.

### Proof of analytic expression for the Gini coefficient

Using the algebraic expression of the Lorenz-curve we can also derive a value for the Gini coefficient defined as

$$Gini = 1 - 2 \int_0^2 L(F) dF = 1 - 2 \left[ \int_0^{0.5} L(F) df + \int_{0.5}^1 L(F) dF \right]. \quad (\text{C.1.15})$$

The overall Lorenz-curve is defined as

$$L(F) = \begin{cases} \frac{1-\alpha}{\alpha} 2^{\frac{1}{\alpha}} F^{\frac{1}{\alpha}+1} & 0 < F < 0.5 \\ 1 - \frac{1+\alpha}{\alpha} 2^{-\frac{1}{\alpha}} (1-F)^{1-\frac{1}{\alpha}} & F \geq 0.5. \end{cases} \quad (\text{C.1.16})$$

With the integration by parts we have

$$Gini = 1 - 1 - 0.5 \left( \frac{\alpha-1}{1+2\alpha} + \frac{1+\alpha}{1-2\alpha} \right) = \frac{3\alpha}{4\alpha^2-1}. \quad \square \quad (\text{C.1.17})$$

A more general result for a non-symmetric Laplace distribution is given in Toda (2012). Note that the Gini coefficient for a double Pareto distribution is always larger than for a standard Pareto (for which it amounts to  $\frac{1}{2\alpha-1}$ ) for a reasonable  $\alpha > 1$ . Note that the symmetric double Pareto



distribution underestimates overall inequality due to not accounting for the close to zero or even negative net worth for low income households. In order to account for the top 1% share of approx. 60% (as being the peak of wealth inequality in the UK) we would require  $\alpha \approx 1.3$ , implying a Gini of approx. 0.68. Recent empirical evidence for more moderate inequality suggest a value of larger than 0.7 for the considered countries (Shorrocks et al., 2016). For  $\alpha \approx 1.3$  the share of the bottom 50% would be given by  $s^{0.5} = 11.5\%$ . For the countries discussed this share is always below 10%.

### Proof of Proposition 6

A change in taxation influences the mean reversion speed  $\mu = -\frac{\tau}{\lambda} - \frac{\gamma^2}{\lambda^2}(1 - \tau)^2 < 0$ . It leads to a jump in  $\mu$ , changing the form of the Pareto-distribution (cf. figure 2(b) of Gabaix et al. (2016) as well as Figure C.5). For an increase in inequality (lower taxes) the Pareto-coefficient decreases firstly at the mode of the distribution only to fade out to the tails. The same holds true for the top-share which, given the average convergence speed  $\phi$ , can be described by

$$s_{t+n}^x = \exp(-\phi n)s_t^x + (1 - \exp(-\phi n))s^x(\tau_t, \infty). \quad (\text{C.1.18})$$

This would be the case for a one-time shock with  $s_{n=0}^x = s_t^x$  and  $s_{n \rightarrow \infty}^x = s^x(\tau_t, \infty)$ . A continuous time representation of the convergence process is given by the following differential equation

$$\frac{ds^x}{dt} = \phi(s^x(\tau_t, \infty) - s^x), \quad (\text{C.1.19})$$

which is solved by the aforementioned equation for an initial condition  $s_t^x$ . The discrete time equivalent would be a mean-reversion process

$$s_t^x - s_{t-1}^x = (1 - \rho)(s^x(\tau_t, \infty) - s_{t-1}^x) \leftrightarrow s_t^x = \rho s_{t-1}^x + (1 - \rho)s^x(\tau_t, \infty), \quad (\text{C.1.20})$$

with  $\Delta t = 1$ . Comparing Equation (C.1.18) with Equation (C.1.20) for  $n = \Delta t = 1$  it is easy to see that  $\rho = \exp(-\phi)$ .

Given the overall time varying nature of the tax rate  $\tau_t$  the long-run steady state  $s^x(\tau_t, \infty)$  is also time varying. Moreover, the reversion speed  $\phi_t$  is also subject to time changes. Thus, the final recursive equation reads

$$s_t^x(\tau_t) = \exp(-\phi_t)s_{t-1}^x + (1 - \exp(-\phi_t))s^x(\tau_t, \infty). \quad \square \quad (\text{C.1.21})$$

## C.2 Robustness with Respect to $\gamma$

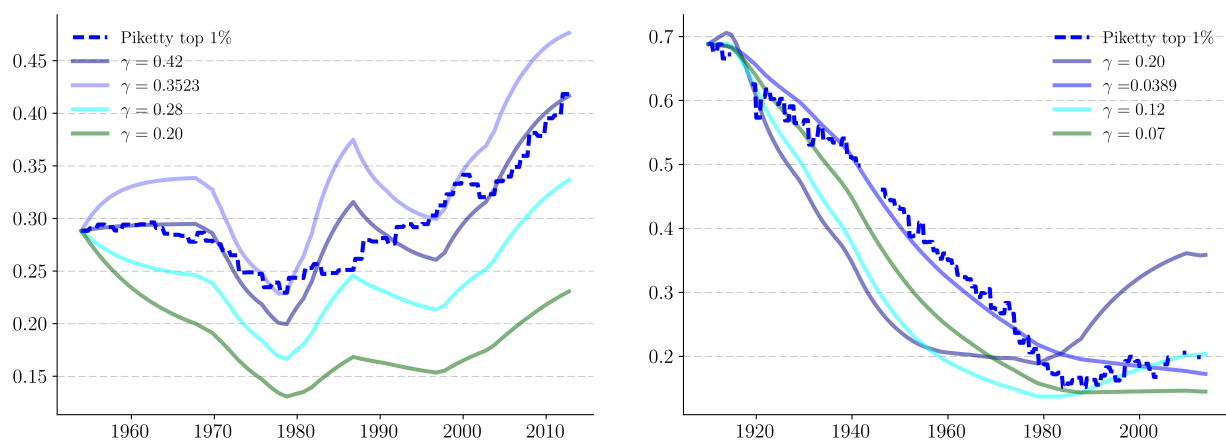


Figure C.1: Variations of  $\gamma$  in the closed-form solutions for the USA (left) and the UK (right).

In Figures C.1 and C.2 we plot the closed-form LOM with different values of  $\gamma$  for all considered countries in order to show that our results are robust even when  $\gamma$  varies considerably. In general - and as formally stated - a lower value of the free parameter  $\gamma$  implies both lower inequality and slower dynamics. This implies a potential effect that while converging to lower equilibrium a low  $\gamma$  can (temporarily) be accompanied by higher inequality (cf. e.g. the case for France with  $\gamma = 0.05$ ). Yet, in the stationary case a lower  $\gamma$  is always accompanied by lower inequality.

Moreover, these simulations imply that in particular for the UK, France and Sweden our results until the 1980s are considerably robust with respect to  $\gamma$ .

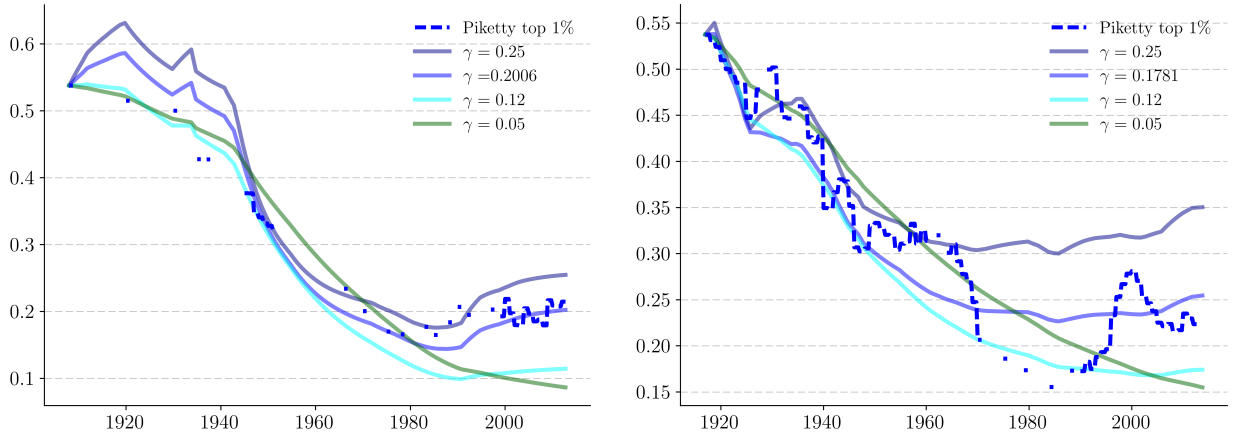


Figure C.2: Variations of  $\gamma$  in the closed-form solutions for Sweden (left) and France (right).

### C.3 Out-of-sample tests for the UK and France

As discussed in the main body correct identification of  $\gamma$  and  $\eta$  for the UK is difficult if one excludes the stable (final) period from the data. Any period before the 1990s is insufficient to nail down  $\gamma$ . Only when considering the 1990s would we be able to conduct a meaningful prediction, but then we would be short of further datapoints to confirm the quality of our predictions. The results are shown in Figure C.3 (left). A further problem is induced by the fact that, for France and Sweden, in a considerable time span only few data points are given.

For the sake of completeness we also perform out-of-sample forecasts for France. While the period from 1950 to 1970 is relatively stable, our above simulations fall short of accounting for the relatively low inequality from 1970 until 1990. Respectively, when excluding the years since 1970 the estimate for  $\gamma$  is around 0.23 regardless of when the calibration period begins. However, the quality of the out-of-sample predictions then crucially depends on the initialization date. If the simulation is then already initialized in the late 1960s, predicted inequality will be higher than documented by the date. Likewise, the level of inequality in the year 2000 is underestimated when we initiate our simulation around 1980. This indicates, again, that the tax changes proceeding the 1980s are very relevant to both forming the level and dynamics of inequality. The predictions are shown in Figure C.3 (right). Since our attempt to match the data for France in the previous

section has led to rather mixed results, it was not to be expected that the model would perform extraordinarily well in the out-of-sample forecasts.

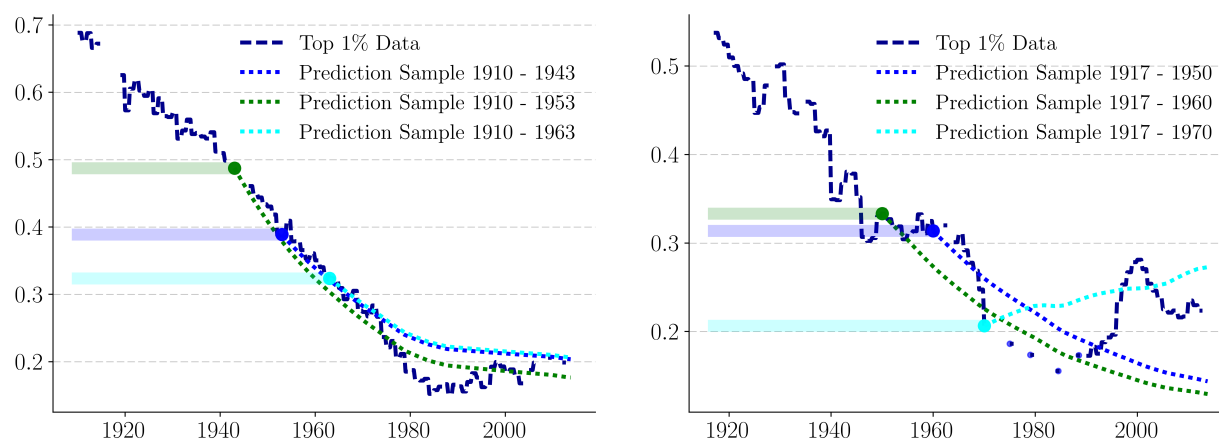


Figure C.3: Out-of-sample predictions for the UK (left) and France (right). Fat bars indicate the respective sample period.

## C.4 Additional Notes on the Stationary Distribution

In Section 4.4 we present a closed-form solution for the stationary distribution of our model. Figure C.5 presents a close-up look at the right tail with a Pareto-fit. The fit misses the extreme tail of the very rich (less than 0.1% of the population). As discussed in Gabaix et al. (2016), this is due to the slow convergence in the tail.<sup>2</sup> The change in inequality - here an increase in inequality - has not yet reached the very right tail.

Figure C.4 presents the histogram of the wealth distribution as created by an exemplary simulation using the USA calibration and tax time series. As predicted by the computation the distribution is symmetric around the mode of 1.

<sup>2</sup>The interested reader is referred in particular to Figure 2 of Gabaix et al. (2016).

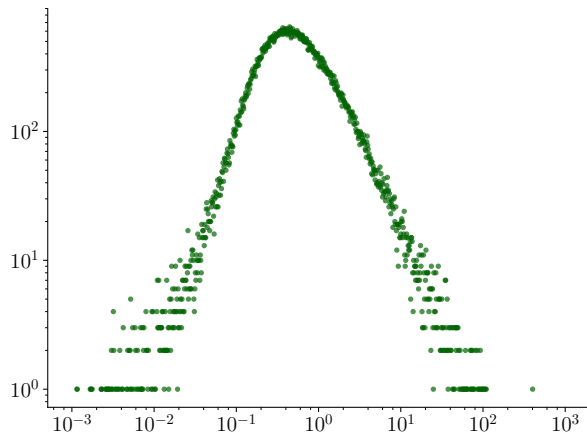


Figure C.4: Log histogram of simulation using US calibration after a very long horizon of 50,000 periods.  $\gamma = 0.3523$ ,  $\tau = 25\%$  and  $N = 100,000$ .

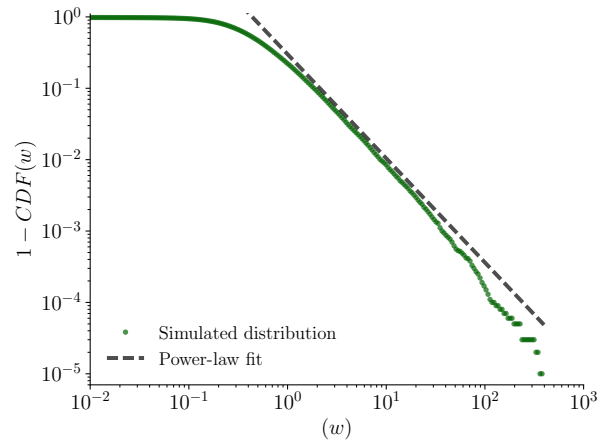


Figure C.5: Right tail of the distribution for the last period for exemplary simulation using the US calibration of  $\gamma = 0.3523$ ,  $\bar{\omega}_{US}^{-1} = 0$  and  $\theta_t = 0.25$  at all  $t$ . The dashed line represents a Pareto-fit with an estimated coefficient of 1.459.



# Bibliography

- ABBATE, A., S. EICKMEIER, W. LEMKE, AND M. MARCELLINO (2016): “The Changing International Transmission of Financial Shocks: Evidence from a Classical Time-Varying FAVAR,” *Journal of Money, Credit and Banking*, 48, 573–601.
- ABEL, A. B. (1990): “Asset prices under habit formation and catching up with the Joneses,” Tech. rep., National Bureau of Economic Research.
- ADJEMIAN, S., H. BASTANI, M. JUILLARD, F. MIHOUBI, G. PERENDIA, M. RATTO, AND S. VILLEMOT (2011): “Dynare: Reference manual, version 4,” Tech. rep., Dynare working papers 1, CEPREMAP.
- AIYAGARI, S. R. (1994): “Uninsured Idiosyncratic Risk and Aggregate Saving,” *The Quarterly Journal of Economics*, 109, 659–84.
- (1995): “Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints, and Constant Discounting,” *Journal of Political Economy*, 103, 1158–75.
- ALSTADSAETER, A., N. JOHANNESSEN, AND G. ZUCMAN (2017): *Tax Evasion and Inequality*, Working Paper.
- ALVAREDO, F., A. B. ATKINSON, AND S. MORELLI (2017): “Top Wealth Shares in the UK Over More than a Century,” Dept. of Economics Research Paper Series No. 01/WP/2017, University Ca’ Foscari of Venice.
- AMIR, R., I. V. EVSTIGNEEV, T. HENS, AND K. R. SCHENK-HOPPÉ (2005): “Market selection and survival of investment strategies,” *Journal of Mathematical Economics*, 41, 105–122.
- ANUFRIEV, M., T. ASSENZA, C. HOMMES, AND D. MASSARO (2008): “Interest rate rules with heterogeneous expectations,” Tech. rep., Universiteit van Amsterdam, Center for Nonlinear Dynamics in Economics and Finance.
- AOKI, S. AND M. NIREI (2017): “Zipf’s Law, Pareto’s Law, and the Evolution of Top Incomes in the U.S.,” *American Economic Journal: Macroeconomics*, forthcoming.
- ASSENMACHER, K. AND S. GERLACH (2008): “Monetary policy, asset prices and macroeconomic conditions: a panel-VAR study,” *National Bank of Belgium Working Paper*.
- ASSENZA, T., P. HEEMEIJER, C. H. HOMMES, AND D. MASSARO (2013): “Individual expectations and aggregate macro behavior.”
- BARRO, R. J. (1990): “The stock market and investment,” *Review of Financial Studies*, 3, 115–131.
- BENHABIB, J., A. BISIN, AND S. ZHU (2011): “The Distribution of Wealth and Fiscal Policy in Economies With Finitely Lived Agents,” *Econometrica*, 79, 123–157.
- (2014): “THE DISTRIBUTION OF WEALTH IN THE BLANCHARD–SYAARI MODEL,” *Macroeconomic Dynamics*, FirstView, 1–16.
- (2015): “The wealth distribution in Bewley economies with capital income risk,” *Journal of Economic Theory*, 159, Part A, 489 – 515.

- BERNANKE, B. AND M. GERTLER (2000): “Monetary policy and asset price volatility,” Tech. rep., National bureau of economic research.
- BERNANKE, B. S., M. GERTLER, AND S. GILCHRIST (1999): “The financial accelerator in a quantitative business cycle framework,” *Handbook of macroeconomics*, 1, 1341–1393.
- BEWLEY, T. (1977): “The permanent income hypothesis: A theoretical formulation,” *Journal of Economic Theory*, 16, 252–292.
- BLANCHARD, O. J. AND C. M. KAHN (1980): “The solution of linear difference models under rational expectations,” *Econometrica: Journal of the Econometric Society*, 1305–1311.
- BLUME, L. AND D. EASLEY (2010): “Heterogeneity, selection, and wealth dynamics,” *Annu. Rev. Econ.*, 2, 425–450.
- BOEHL, G. (2017): *On the Evolutionary Fitness of Rational Expectations*, *In preparation*.
- BORIO, C. E. AND P. W. LOWE (2002): *Asset prices, financial and monetary stability: exploring the nexus*, BIS working paper.
- BRANCH, W. A. (2004): “The Theory of Rationally Heterogeneous Expectations: Evidence from Survey Data on Inflation Expectations\*,” *The Economic Journal*, 114, 592–621.
- BRANCH, W. A. AND B. MCGOUGH (2009): “A New Keynesian model with heterogeneous expectations,” *Journal of Economic Dynamics and Control*, 33, 1036–1051.
- (2010): “Dynamic predictor selection in a new Keynesian model with heterogeneous expectations,” *Journal of Economic Dynamics and Control*, 34, 1492–1508.
- BRAZIER, A., R. HARRISON, M. KING, AND T. YATES (2008): “The danger of inflating expectations of macroeconomic stability: heuristic switching in an overlapping-generations monetary model,” *International Journal of Central Banking*, 4, 219–254.
- BROCK, W. A. AND C. H. HOMMES (1997): “A rational route to randomness,” *Econometrica*, 1059–1095.
- (1998): “Heterogeneous beliefs and routes to chaos in a simple asset pricing model,” *Journal of Economic dynamics and Control*, 22, 1235–1274.
- BROCK, W. A., C. H. HOMMES, AND F. O. O. WAGENER (2009): “More hedging instruments may destabilize markets,” *Journal of Economic Dynamics and Control*, 33, 1912–1928.
- CAGETTI, M. AND M. D. NARDI (2009a): “Estate Taxation, Entrepreneurship, and Wealth,” *American Economic Review*, 99, 85–111.
- (2009b): “Estate Taxation, Entrepreneurship, and Wealth,” *American Economic Review*, 99, 85–111.
- CAMPBELL, J. Y. (1999): “Asset prices, consumption, and the business cycle,” *Handbook of macroeconomics*, 1, 1231–1303.
- CAMPBELL, J. Y. AND L. M. VICEIRA (2002): *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors*, no. 9780198296942 in OUP Catalogue, Oxford University Press.
- CASTANEDA, A., J. DIAZ-GIMENEZ, AND J.-V. RIOS-RULL (2003): “Accounting for the U.S. Earnings and Wealth Inequality,” *Journal of Political Economy*, 111, 818–857.
- CECCHETTI, S. G. (2000): *Asset prices and central bank policy*, Centre for Economic Policy Research.
- CLEMENS, C. AND M. HEINEMANN (2015): “Endogenous growth and wealth inequality under incomplete markets and idiosyncratic risk,” *Journal of Macroeconomics*, 45, 300–317.



- CONLISK, J. (1996): “Why bounded rationality?” *Journal of economic literature*, 34, 669–700.
- COOTNER, P. H. (1964): “The random character of stock market prices,” .
- CRAIG, C. C. (1936): “On the Frequency Function of  $xy$ ,” *Ann. Math. Statist.*, 7, 1–15.
- DE GRAUWE, P. (2011): “Animal spirits and monetary policy,” *Economic theory*, 47, 423–457.
- DE GRAUWE, P. AND C. MACCHIARELLI (2013): *Animal spirits and credit cycles*, CESifo Working Paper Series.
- DE LONG, J. B., A. SHLEIFER, L. H. SUMMERS, AND R. J. WALDMANN (1990): “Positive feedback investment strategies and destabilizing rational speculation,” *the Journal of Finance*, 45, 379–395.
- DOMEIJ, D. AND J. HEATHCOTE (2004): “On The Distributional Effects Of Reducing Capital Taxes,” *International Economic Review*, 45, 523–554.
- DU RIETZ, G. AND M. HENREKSON (2014): “Swedish Wealth Taxation (1911–2007),” Working Paper Series 1000, Research Institute of Industrial Economics.
- EPSTEIN, L. G. AND S. E. ZIN (1989): “Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework,” *Econometrica: Journal of the Econometric Society*, 937–969.
- EVANS, G. W. AND S. HONKAPOHJA (2003): “Expectations and the stability problem for optimal monetary policies,” *The Review of Economic Studies*, 70, 807–824.
- EVANS, G. W., S. HONKAPOHJA, AND R. MARIMON (2001): “Convergence in monetary inflation models with heterogeneous learning rules,” *Macroeconomic Dynamics*, 5, 1–31.
- EVSTIGNEEV, I. V., T. HENS, AND K. R. SCHENK-HOPPÉ (2002): “Market selection of financial trading strategies: Global stability,” *Mathematical Finance*, 12, 329–339.
- FAMA, E. F. (1965): “The Behavior of Stock-Market Prices,” *The Journal of Business*, 38, 34–105.
- FAMA, E. F. AND K. R. FRENCH (1988): “Dividend yields and expected stock returns,” *Journal of financial economics*, 22, 3–25.
- FERNHOLZ, R. AND R. FERNHOLZ (2014): “Instability and concentration in the distribution of wealth,” *Journal of Economic Dynamics and Control*, 44, 251–269.
- FRIEDMAN, M. (1953): *Essays in positive economics*, Chicago: University of Chicago Press.
- GABAIX, X., J.-M. LASRY, P.-L. LIONS, AND B. MOLL (2016): “The Dynamics of Inequality,” *Econometrica*, 84, 2071–2111.
- GALÍ, J. (2008): *Introduction to Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework*, Princeton University Press.
- (2013): “Monetary policy and rational asset price bubbles,” Tech. rep., National Bureau of Economic Research.
- GARBINTI, B., J. GOUPILLE-LEBRET, AND T. PIKETTY (2017): “Accounting for Wealth Inequality Dynamics: Methods, Estimates and Simulations for France (1800-2014),” CEPR Discussion Papers 11848, C.E.P.R. Discussion Papers.
- GERTLER, M., N. KIYOTAKI, ET AL. (2010): “Financial intermediation and credit policy in business cycle analysis,” *Handbook of monetary economics*, 3, 547–599.
- GREENWOOD, R. AND A. SHLEIFER (2014a): “Expectations of returns and expected returns,” *Review of Financial Studies*, 27, 714–746.

- (2014b): “Expectations of Returns and Expected Returns,” *Review of Financial Studies*, 27, 714–746.
- HENS, T. AND K. R. SCHENK-HOPPÉ (2005): “Evolutionary stability of portfolio rules in incomplete markets,” *Journal of mathematical economics*, 41, 43–66.
- HOMMES, C. (2011): “The heterogeneous expectations hypothesis: Some evidence from the lab,” *Journal of Economic dynamics and control*, 35, 1–24.
- (2013): *Behavioral rationality and heterogeneous expectations in complex economic systems*, Cambridge University Press.
- HOMMES, C., J. SONNEMANS, J. TUINSTRA, AND H. VAN DE VELDEN (2005): “A strategy experiment in dynamic asset pricing,” *Journal of Economic Dynamics and Control*, 29, 823–843.
- HOMMES, C. H. (2006): “Heterogeneous agent models in economics and finance,” *Handbook of computational economics*, 2, 1109–1186.
- HUBMER, J., P. KRUSELL, AND J. ANTHONY A. SMITH (2016): “The Historical Evolution of the Wealth Distribution: A Quantitative-Theoretic Investigation,” NBER Working Papers 23011, National Bureau of Economic Research, Inc.
- HUGGETT, M. (1993): “The risk-free rate in heterogeneous-agent incomplete-insurance economies,” *Journal of Economic Dynamics and Control*, 17, 953–969.
- IMF (2003): Washington, International Monetary Fund.
- JUDD, K. L. (1985): “Redistributive taxation in a simple perfect foresight model,” *Journal of Public Economics*, 28, 59–83.
- (1998): *Numerical methods in economics*, MIT press.
- KAPLAN, G., B. MOLL, AND G. L. VIOLANTE (2016): “Monetary Policy According to HANK,” NBER Working Papers 21897, National Bureau of Economic Research, Inc.
- KARLIN, S. AND H. TAYLOR (1981): *A Second Course in Stochastic Processes*, 2, Academic Press.
- KAYMAK, B. AND M. POSCHKE (2016): “The evolution of wealth inequality over half a century: The role of taxes, transfers and technology,” *Journal of Monetary Economics*, 77, 1–25.
- KEYNES, J. M. (1936): *General theory of employment, interest and money*, Atlantic Publishers & Dist.
- KIRMAN, A. (1992): “Whom or what does the representative individual represent?” *The Journal of Economic Perspectives*, 6, 117–136.
- KLIEM, M. AND H. UHLIG (2016): “Bayesian estimation of a dynamic stochastic general equilibrium model with asset prices,” *Quantitative Economics*, 7, 257–287.
- KRUSELL, P. AND T. SMITH (1998): “Income and Wealth Heterogeneity in the Macroeconomy,” *Journal of Political Economy*, 106, 867–896.
- LANSING, K. J. (1999): “Optimal redistributive capital taxation in a neoclassical growth model,” *Journal of Public Economics*, 73, 423–453.
- LEVHARI, D. AND T. N. SRINIVASAN (1969): “Optimal savings under uncertainty,” *The Review of Economic Studies*, 36, 153–163.
- LJUNGQVIST, L. AND T. J. SARGENT (2012): *Recursive macroeconomic theory*, MIT press.

- LJUNGQVIST, L. AND H. UHLIG (2000): "Tax policy and aggregate demand management under catching up with the Joneses," *American Economic Review*, 356–366.
- (2015): "Comment on the Campbell-Cochrane habit model," *Journal of Political Economy*, 123, 1201–1213.
- LUCAS, R. E. (1972): "Expectations and the Neutrality of Money," *Journal of economic theory*, 4, 103–124.
- LUNDBERG, J. AND D. WALDENSTRÖM (2017): "Wealth Inequality in Sweden: What can we Learn from Capitalized Income Tax Data?" *Review of Income and Wealth*, n/a–n/a.
- MANKIW, N. G., R. REIS, AND J. WOLFERS (2003): "Disagreement about inflation expectations," Tech. rep., National Bureau of Economic Research.
- MARTIN, A. AND J. VENTURA (2010): "Economic growth with bubbles," Tech. rep., National Bureau of Economic Research.
- (2011): "Theoretical notes on bubbles and the current crisis," *IMF Economic Review*, 59, 6–40.
- MIAO, J., P. WANG, AND L. XU (2016): "Stock market bubbles and unemployment," *Economic Theory*, 61, 273–307.
- MIAO, J., P. WANG, AND Z. XU (2012): "A Bayesian DSGE model of stock market bubbles and business cycles," Tech. rep., working paper, Boston University.
- MIRANDA, M. J. AND P. L. FACKLER (2004): *Applied computational economics and finance*, MIT press.
- MUTH, J. F. (1961): "Rational Expectations and the Theory of Price Movements," *Econometrica*, 29, 315–335.
- PFAJFAR, D. AND E. SANTORO (2008): *Asymmetries in inflation expectation formation across demographic groups*, University of Cambridge, Faculty of Economics.
- (2010): "Heterogeneity, learning and information stickiness in inflation expectations," *Journal of Economic Behavior & Organization*, 75, 426–444.
- PFAJFAR, D. AND B. ZAKELJ (2012): "Uncertainty and disagreement in forecasting inflation: Evidence from the laboratory," *European Banking Center Discussion Paper*.
- PIKETTY, T. (2014): *Capital in the twenty-first century*, Boston: Harvard University Press.
- PULLEY, L. B. (1983): "Mean-variance approximations to expected logarithmic utility," *Operations Research*, 31, 685–696.
- QUADRINI, V. (2000): "Entrepreneurship, Saving and Social Mobility," *Review of Economic Dynamics*, 3, 1–40.
- ROINE, J. AND D. WALDENSTRÖM (2015): "Chapter 7 - Long-Run Trends in the Distribution of Income and Wealth," in *Handbook of Income Distribution*, ed. by A. B. Atkinson and F. Bourguignon, Elsevier, vol. 2 of *Handbook of Income Distribution*, 469 – 592.
- RUDEBUSCH, G. D. AND E. T. SWANSON (2012): "The bond premium in a dsge model with long-run real and nominal," *American Economic Journal: Macroeconomics*, 4, 105–143.
- SAEZ, E. AND G. ZUCMAN (2016): "Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data," *The Quarterly Journal of Economics*, 131, 519–578.
- SARGENT, T. (2008): "Rational expectations, The Concise Encyclopedia of Economics," .
- SHILLER, R. J. (1981): "The use of volatility measures in assessing market efficiency," *The Journal of Finance*, 36, 291–304.

- SHORROCKS, A., J. B. DAVIES, AND R. LLUBERAS (2016): "Global Wealth Databook 2016," Research institute, Credit Suisse.
- SIMON, H. A. (1955): "A behavioral model of rational choice," *The quarterly journal of economics*, 69, 99–118.
- SIMS, C. A. (1980): "Macroeconomics and reality," *Econometrica: Journal of the Econometric Society*, 1–48.
- SMETS, F. AND R. WOUTERS (2003): "An estimated dynamic stochastic general equilibrium model of the euro area," *Journal of the European economic association*, 1, 1123–1175.
- STIGLITZ, J. E. (1969): "The Effects of Income, Wealth, and Capital Gains Taxation on Risk-Taking," *The Quarterly Journal of Economics*, 83, 263–283.
- STRAUB, L. AND I. WERNING (2014): "Positive Long Run Capital Taxation: Chamley-Judd Revisited," NBER Working Papers 20441, National Bureau of Economic Research, Inc.
- TALLARINI, T. D. (2000): "Risk-sensitive real business cycles," *Journal of monetary Economics*, 45, 507–532.
- TODA, A. A. (2012): "The double power law in income distribution: Explanations and evidence," *Journal of Economic Behavior & Organization*, 84, 364 – 381.
- TUINSTRAL, J. AND F. O. WAGENER (2007): "On learning equilibria," *Economic Theory*, 30, 493–513.
- WINKLER, F. (2014): *The Role of Learning for Asset Prices, Business Cycles and Monetary Policy*.
- WOLD, H. O. A. AND P. WHITTLE (1957): "A Model Explaining the Pareto Distribution of Wealth," *Econometrica*, 25, 591–595.
- WOODFORD, M. (2003): *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton University.