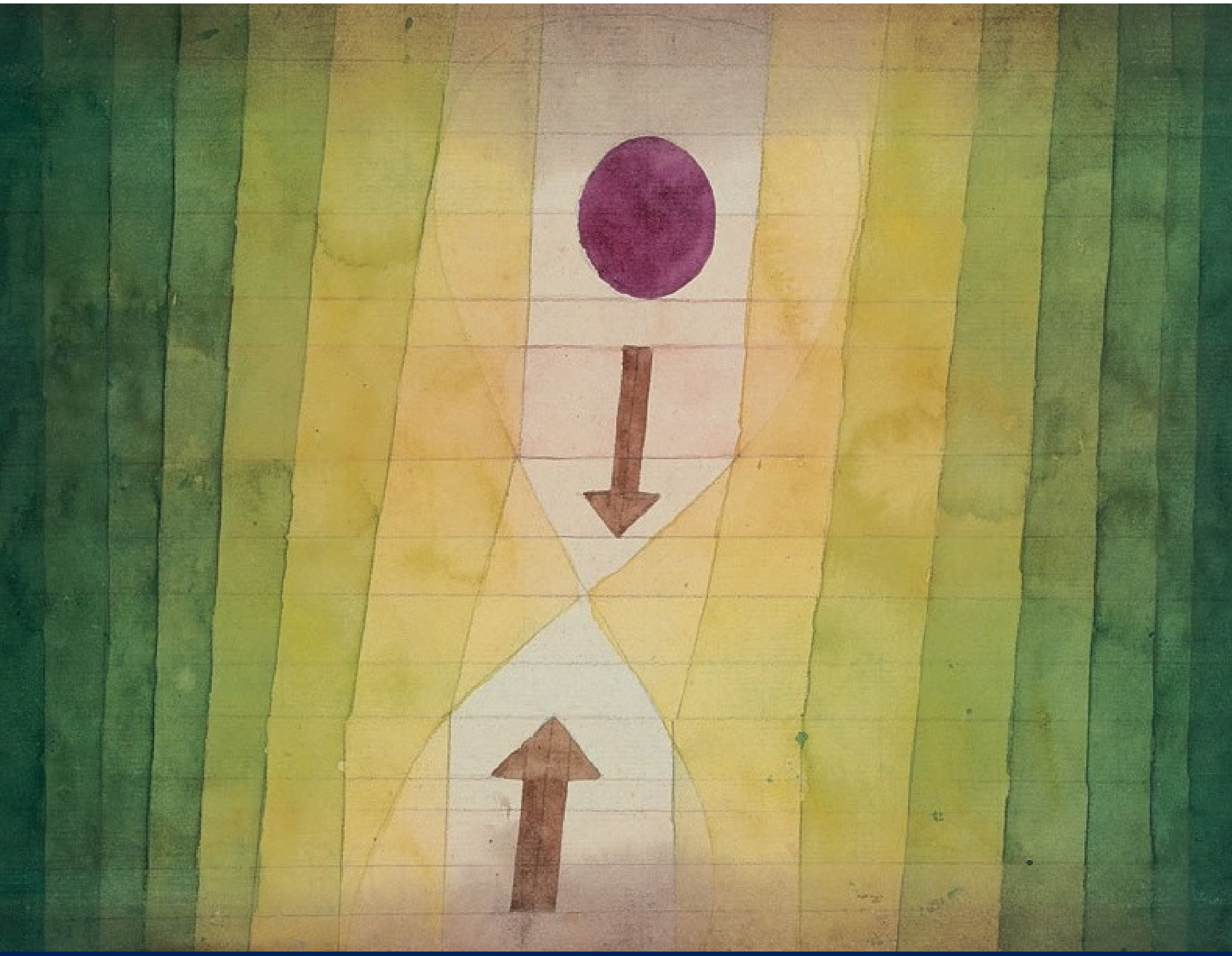
Institute for International Economic Studies
Monograph Series No. 124

Stockholm University

Essays on Income Risk, Portfolio Choices and the Macroeconomy

Gualtiero Azzalini



Doctoral Thesis in Economics at Stockholm University, Sweden 2023

Essays on Income Risk, Portfolio Choices and the Macroeconomy

Gualtiero Azzalini

Academic dissertation for the Degree of Doctor of Philosophy in Economics at Stockholm University to be publicly defended on Wednesday 14 June 2023 at 13.00 in Nordenskiöldssalen, Geovetenskapens hus, Svante Arrhenius väg 12.

Abstract

Business cycle asymmetry of earnings pass-through

How does the firm's role as an insurance provider vary over the business cycle? Using Swedish administrative data, I document that idiosyncratic firm productivity shocks are passed through workers' earnings asymmetrically. In non-recessions, firms are good insurers against negative shocks. In downturns, they pass through a larger share of their shock. Regardless of the state of the economy, instead, positive shocks are mainly passed through when sizeable. I rationalize these findings using a directed search model of the labor market with recursive contracts. Moral hazard risk associated with on-the-job search is key to generating pass-through and the increased risk of firm disaster in recessions is necessary for matching the empirical facts. As the wage growth distribution features procyclical skewness and acyclical variance, the model also suggests a new mechanism for explaining trends in income risk variation over the business cycle. Welfare calculations reveal that workers would be willing to give up a non-negligible share of consumption to avoid this source of uncertainty.

Inferring income properties from portfolio choices

Two main views exist on the nature of the labor income process: according to one, income shocks are very persistent and agents face similar life-cycle profiles - *Restricted Income Profiles* (RIP); according to the other, income shocks are not very persistent and life-cycle profiles are individual-specific - *Heterogeneous Income Profiles* (HIP). This paper studies the implications of these two views in a portfolio choice model in order to discover identification restrictions allowing to discern between them. I find that HIP and RIP imply different life-cycle patterns of the participation and conditional risky share choices but similar patterns of consumption and saving. Crucial for this result is the inclusion of cyclical skewness in the stochastic process for income, which enables us to correctly estimate the part of income risk deriving from the persistence of the shocks.

Preference heterogeneity and portfolio choices over the wealth distribution

What are the key elements required in generating portfolio choices over the wealth distribution in line with the data? In this paper, we argue that capturing preference heterogeneity across individuals is one of them. Using a partial equilibrium Bewley-type model with endogenous portfolio choice and cyclical skewness in labor income shocks, we show that heterogeneity in risk aversion, impatience and portfolio diversification is crucial to match the empirical schedules of unconditional risky share, participation and share of idiosyncratic variance in individual portfolios. At the same time, these elements generate dispersion in wealth through their heterogeneous effects on individuals' investment decisions resulting in a cross-sectional wealth distribution that provides a close fit of the data, particularly at the very top.

Keywords: *Macroeconomics, household finance, income risk, portfolio choice, wealth inequality, heterogeneous agents, insurance, search and matching.*

Stockholm 2023
http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-216752

ISBN 978-91-8014-368-4
ISBN 978-91-8014-369-1
ISSN 0346-6892

Stockholm University

Department of Economics

Stockholm University, 106 91 Stockholm

ESSAYS ON INCOME RISK, PORTFOLIO CHOICES AND THE MACROECONOMY

Gualtiero Azzalini

Stockholm University

Essays on Income Risk, Portfolio Choices and the Macroeconomy

Gualtiero Azzalini

ISBN print 978-91-8014-368-4
ISBN PDF 978-91-8014-369-1
ISSN 0346-6892

Cover photo: Paul Klee, Vor dem Blitz, 1923 (Cropped). WikiArt - Visual Art Encyclopedia
https://www.wikiart.org/en/paul-klee/before-the-blitz-1923.

Back-cover photo by Hanna Weitz, Stockholm University.

Printed in Sweden by Universitetsservice US-AB, Stockholm 2023

Doctoral Dissertation
Department of Economics
Stockholm University

Abstracts

Business cycle asymmetry of earnings pass-through

How does the firm's role as an insurance provider vary over the business cycle? Using Swedish administrative data, I document that idiosyncratic firm productivity shocks are passed through workers' earnings asymmetrically. In non-recessions, firms are good insurers against negative shocks. In downturns, they pass through a larger share of their shock. Regardless of the state of the economy, instead, positive shocks are mainly passed through when sizeable. I rationalize these findings using a directed search model of the labor market with recursive contracts. Moral hazard risk associated with on-the-job search is key to generating pass-through and the increased risk of firm disaster in recessions is necessary for matching the empirical facts. As the wage growth distribution features procyclical skewness and acyclical variance, the model also suggests a new mechanism for explaining trends in income risk variation over the business cycle. Welfare calculations reveal that workers would be willing to give up a non-negligible share of consumption to avoid this source of uncertainty.

Inferring income properties from portfolio choices

Two main views exist on the nature of the labor income process: according to one, income shocks are very persistent and agents face similar life-cycle profiles - *Restricted Income Profiles* (RIP); according to the other, income shocks are not very persistent and life-cycle profiles are individual-specific - *Heterogeneous Income Profiles* (HIP). This paper studies the implications of these two views in a portfolio choice model in order to discover identification restrictions allowing to discern between them. I find that HIP and RIP imply different life-cycle

patterns of the participation and conditional risky share choices but similar patterns of consumption and saving. Crucial for this result is the inclusion of cyclical skewness in the stochastic process for income, which enables us to correctly estimate the part of income risk deriving from the persistence of the shocks.

Preference heterogeneity and portfolio choices over the wealth distribution *(with Markus Kondziella and Zoltán Rácz)*
What are the key elements required in generating portfolio choices over the wealth distribution in line with the data? In this paper, we argue that capturing preference heterogeneity across individuals is one of them. Using a partial equilibrium Bewley-type model with endogenous portfolio choice and cyclical skewness in labor income shocks, we show that heterogeneity in risk aversion, impatience and portfolio diversification is crucial to match the empirical schedules of unconditional risky share, participation and share of idiosyncratic variance in individual portfolios. At the same time, these elements generate dispersion in wealth through their heterogeneous effects on individuals' investment decisions resulting in a cross-sectional wealth distribution that provides a close fit of the data, particularly at the very top.

Acknowledgements

I started a Ph.D. moved by a genuine desire to understand economic phenomena with rigorous analytical tools. If today this curiosity has been partially satiated is thanks to the people I met along the way.

I am deeply indebted to Per Krusell, Kurt Mitman, Kieran Larkin and Paolo Sodini. The impact they had on me exceeds by far the content of this thesis.

I am grateful to Per for constantly pushing me to challenge my work, think about the big picture, formulate precise arguments and build up self-confidence. I thank Kurt for always pointing out a solution and for his unwavering encouragement, especially during stressful moments. I owe Kieran for guiding me through the practical implementation of my ideas and nudging me to extract more from them than I would have done myself. I am thankful to Paolo for spurring my interest in the household finance literature and for always being curious and enthusiastic about my research.

I also wish to thank Timo Boppart and Joshua Weiss for their comments and advices - particularly before the job market - and to acknowledge Arash Nekoei, Mitch Downey, John Hassler, Peter Nilsson and José Montalbán Castilla for providing valuable feedback.

During the Ph.D. I was lucky to work with two exceptional coauthors, Markus Kondziella and Zoltán Rácz. I have learned a great deal from them and I am glad that our professional relationships have transformed into solid friendships. I found two other good friends in José-Elías Gallegos and Ida Kristine Haavi, who have also been excellent office mates.

I am also thankful to Magnus Åhl, Mattias Almgren, Kasper Kragh Balke, Agneta Berge, Iacopo Bianchi, Max Bres, Cristina Clerici, Mohammadreza Farajpour, Richard Foltyn, Stefan Hinkelmann, Philipp Hochmuth, Markus Karlman, Karin Kinnerud, John Kramer, Carolina Lindholm, Evelina Linnros, Juan Llavador Peralt, Francesco Loiacono, Jonna Olsson, Fredrik Paues, Markus Peters, Sreyashi Sen, Xueping

Sun, Sebastian Tebbe, Claire Thürwächter, Has van Vlokhoven and Tillmann von Carnap-Bornheim, who have all contributed - in different ways - to the realization of this work.

A special thanks also goes to Christina Lönnblad, Ulrika Gålnander, Tove Lundström, Hanna Weitz and Karl Eriksson. The fact that during these years I could solely focus on my research was all thanks to their help, efficiency and kindness.

Embarking on the route of academic research would have not been possible without the guidance of Carlo Favero, Francesco Giavazzi and Alberto Alesina. I am profoundly grateful to them for being invaluable mentors since my years at Bocconi.

I also wish to thank my family for their unconditional support. I am particularly thankful to my parents, Marzia and Maurizio, and to my sister Vera for always backing me up and encouraging me to do my best. Special acknowledgment also goes to my uncle Adalberto, for having directed me towards this professional path, and to my cousin Andrea, for always being a guide in this journey.

This dissertation is dedicated to Vanessa. Everything I have accomplished so far is due to her determination to make me every day a better version of myself.

Gualtiero Azzalini
Stockholm, Sweden
May 2023

Contents

Introduction

This thesis consists of three self-contained essays. All of them combine the usage of micro and macro data and quantitative models to study how agents balance their exposure to income risk when facing idiosyncratic and aggregate shocks.

The first chapter, *"Business cycle asymmetry of earnings pass-through"*, analyzes how endogenous income risk emerges from the optimal risk-sharing allocation between workers and firms.

While understanding how much firms insure their workers against salary fluctuations is a long-standing topic in economics, this paper examines this question from a new angle by investigating how firms' ability to do that varies over the business cycle. As labor market and financial frictions bind differently in booms and recessions, there are reasons to suspect that shocks' transmission to workers' compensations might vary with aggregate shocks.

Using Swedish administrative data, I document that the pass-through of idiosyncratic firm shocks to workers' earnings is indeed asymmetric over the cycle. Firms insure workers against negative shocks in non-recessionary periods, but they do much less so in downturns. Positive shocks, on the other hand, are shared with employees especially if sizable, and this holds regardless of the state of the economy.

I further show that these empirical patterns can be rationalized using a directed search model of the labor market with on-the-job search, risk-averse workers, and firm commitment. The key element in the model is a trade-off that firms face when choosing the terms of the employment relationship with their workers. On the one hand, insuring risk-averse workers against wage fluctuations enables firms to avoid paying the volatility risk premium implied by the concavity of the workers' utility function. On the other, guaranteeing the employees stable compensation does not allow firms to align the workers' search incentives with the firms' own desire to sustain the match or not.

In addition to matching earnings pass-through asymmetries found in the data, the model provides a new explanation for the business cycle trends in income risk documented by a recent empirical literature. More in detail, since the pass-through of negative shocks is on average larger in recessions and that of positive ones is acyclical, the model-generated wage growth distribution features procyclical skewness. Finally, I evaluate the welfare cost of business cycles and find that they are substantial in this framework.

The second chapter, *"Inferring income properties from portfolio choices"*, shows that endogenous income risk coming from agents' portfolio choices is informative on the true nature of the labor income process.

Although the literature trying to understand the properties of the income process is vast, two main hypotheses have emerged: according to one, income shocks are very persistent and agents face similar life-cycle profiles - Restricted Income Profiles (RIP); according to the other, income shocks are not very persistent and life-cycle profiles are individual-specific - Heterogeneous Income Profiles (HIP).

In this paper, I study whether agents' portfolio choices contain relevant information allowing us to discern which of the two views is more supported by the data. The main idea is that, since diverse types of income risk imply different portfolio allocation decisions, by looking at the latter the researcher can infer properties of the income process.

Because taking into account the cyclical skewness of income shocks results in similar estimates for the shocks' persistence for both HIP and RIP, I find that the profiles of the mean and variance of consumption over the life cycle are very much alike and, thus, do not have a strong identification power.

However, HIP and RIP imply different average life-cycle profiles for stock market participation and for the choice of the conditional risky share. Due to the effect of cyclical skewness on the riskiness of human capital, the HIP process implies much less heterogeneity in participation rates across people with different average income growth rates and a "butterfly pattern" for the conditional risky share. The latter means that agents with higher average growth rates have a lower conditional risky share at young ages compared to individuals with low growth rates, and they catch up at around forty years old when the order of this pattern is reversed.

Comparing the model-generated profiles and their empirical equivalents using Swedish administrative data, I find that the latter provides slightly stronger support for the RIP than the HIP hypothesis.

In the third chapter, *"Preference heterogeneity and portfolio choices over the wealth distribution"*, jointly written with Markus Kondziella and Zoltán Rácz, we show that endogenous income risk ensuing from preference heterogeneity across individuals helps explain wealth inequality.

Starting from the fact that explaining individuals' portfolio choices and cross-sectional wealth inequality remain two challenging issues in household finance and macroeconomics, we show that connecting the two literatures can help to address both issues simultaneously.

Specifically, we add to the standard incomplete markets macro model endogenous portfolio choice, a non-normal return process, cyclical skewness in labor income shocks, Epstein-Zin preferences and preference heterogeneity including heterogeneity in individuals' time preference rate, risk aversion and ability or inclination towards portfolio diversification.

Estimating the model parameters governing the heterogeneity in preferences to match the increasing risky share, participation rate and share of idiosyncratic return risk over the wealth distribution documented in the data, we find that to get a good match the economy has to be populated by two types of agents: one featuring higher risk aversion and impatience and the other characterized by lower values of the same parameters. Indeed, as agents of the latter type endogenously end up at the top of the distribution, the model is able to capture the increasing relation between wealth and risky share found empirically and, in turn, to generate wealth inequality at the top.

Examining the results of our benchmark specification with counterfactual economies in which we shut down different components one at a time, we find that the model fit is always worsened.

Chapter 1

Business cycle asymmetry of earnings pass-through

I am grateful to Per Krusell, Kurt Mitman, Kieran Larkin, Paolo Sodini, Timo Boppart and Joshua Weiss for advice, discussions and support. I also wish to thank Agneta Berge, Tobias Broer, Richard Foltyn, José-Elías Gallegos, John Hassler, Stefan Hinkelmann, Philipp Hochmuth, Markus Kondziella, John Kramer, Arash Nekoei, Zoltán Rácz, Claire Thürwächter and participants at the IIES Macro Group for helpful feedback and comments. I am thankful to IFAU for access to and help with the Swedish administrative data.

1.1 Introduction

The extent to which firms insure their workers against salary fluctuations is a long-standing question in economics (Knight, 1921), and it is based on the idea that an unequal allocation of the match's surplus uncertainty between employers and employees is desirable if they have different attitudes towards risk (Baily, 1974; Azariadis, 1975).

Increased availability of high-quality matched employer-employee data has led to greater scrutiny of this issue. In particular, recent empirical papers have found that firms shield well workers' compensations against idiosyncratic productivity shocks they are exposed to[1], but also that the degree of insurance is not full, especially if the shocks are persistent (for a review, see Guiso and Pistaferri, 2020).

This literature has, however, predominantly focused on studying this phenomenon unconditionally of the state of the economy. There are reasons to suspect that firms' ability to insure workers might vary with aggregate shocks as, for instance, labor market and financial frictions bind differently over the business cycle. Considering that income risk is a crucial driver of individuals' decisions, taking this additional source of uncertainty into account is essential for the design of welfare-improving policies.[2]

This paper fills this gap by examining the heterogeneity in the transmission of idiosyncratic firm productivity shocks to workers' salaries over the business cycle. Specifically, I document new facts on earnings pass-through asymmetries using Swedish registry data and build a search model of the labor market to replicate these patterns, finding a key role for the increased risk of firm disaster in recessions.

More in detail, using matched employer-employee data covering

[1] Among the shocks against which a firm can insure its workers, idiosyncratic ones stand out because they can be diversified via financial markets (Pagano, 2020). The idea is that the stock market should not price idiosyncratic risk and, therefore, a firm providing insurance against them should not bear a higher cost of equity.

[2] Several papers have analyzed the effect of income risk related to the pass-through of firm shocks on workers' portfolio choices, precautionary savings, and insurance demand. For a comprehensive review, see Guiso and Pistaferri (2020).

the universe of non-financial private companies, employment relationships, and workers in Sweden for the period 2004-2018, I document the following facts. First, earnings pass-through of idiosyncratic firm shocks is asymmetric in the state of the business cycle.[3] Firms insure workers against negative shocks in non-recessionary periods, but they do much less so in downturns. Positive shocks, on the other hand, are shared with employees especially if sizable, and this holds regardless of the state of the economy. Compared to non-recessions, the earnings' elasticity to negative shocks in recessions is more than three times higher (0.020 *vs.* 0.006), even in the most conservative specification. In monetary terms, this means a reduction of approximately 1,500 SEK (in 2020 terms) in the annual salary of a worker employed at a firm experiencing a one standard deviation shock. Given that in recessions the total employment share of firms exposed to more negative shocks is around 15 percent, the aggregate implications of this phenomenon are non-negligible.

Second, negative idiosyncratic productivity shocks are on average more adverse events for the firm if they occur in downturns. In particular, the share of firms experiencing mass layoffs upon receiving these shocks in recession - defined, following von Wachter et al. (2009), as firms whose workforce shrinks by more than 30 percent in two years - is approximately 7 percentage points higher compared to the same quantity in normal times. Combined with the fact that it takes around three years from the shock for the average firm to return to positive employment growth, this indicates that negative idiosyncratic shocks have larger and more persistent effects on firms' profitability in recessions.

On the theoretical side, I rationalize these empirical patterns using a directed search model of the labor market similar to the ones used by Menzio and Shi (2010) and Balke and Lamadon (2022). Specifically, I extend the mechanism of the latter paper to an economy with busi-

[3]Data on hours worked are available for a limited time period and sample of firms. For this reason, I focus on earnings and from here onward I will refer to earnings and wages interchangeably.

ness cycles. The block recursivity property of the framework allows tractability in an environment with aggregate uncertainty.

In this setting, the key elements to generating realistic pass-through are on-the-job search, workers' risk aversion, and firm commitment. In each period, currently matched firms need to deliver the worker a certain amount of utility by selecting the current wage and state-contingent contracts for the next period formulated as promised utilities which, in turn, affect on-the-job search choices. This decision entails a trade-off. On the one hand, the firm wants to insure risk-averse workers against wage fluctuations to avoid paying the volatility risk premium implied by the concavity of the utility function. On the other, not adjusting the terms of the employment contract results in waiving the opportunity to align the worker's search incentives with the firm's desire to keep the worker and sustain or end the match.

Consider a firm exposed to a persistent negative productivity shock. Since the shock reduces the expected value of the match, the firm has a lower incentive to continue the employment relationship, so it optimally diminishes the worker's share of the surplus by promising her lower utility, until this choice is balanced with the cost for the additional risk premium to be paid. This action has two consequences. First, the job-finding rate is a decreasing function of utility, so the probability that the worker is poached increases. Second, as lower promised utilities can be sustained with a diminished stream of consumption, wages are cut. The magnitude of this mechanism is amplified by a higher persistence of the shocks or a lower cost of deviating from full insurance.

In addition to the forces generating the shocks' pass-through, two key elements enable the model to match the asymmetries found in the data. The first is modeling that negative idiosyncratic firm shocks are more adverse events in recessions, which is needed to replicate the larger pass-through of these shocks in downturns. While the general equilibrium effect from a lower job finding probability in recessions pushes the framework toward the desired direction by requiring a more significant decrease in promised utility to achieve the same probability that the worker leaves, I find that it is quantitatively not strong enough to generate the asymmetries found in the data. Thus, I introduce an additional disaster state in which, upon entering, firms experience low productivity for a protracted period of time, and which I discipline using the empirical evidence on mass layoffs. As explained above, the higher the persistence of the shocks, the stronger the mechanism. Therefore, if a negative shock has a larger and more persistent effect on the expected value of the match in recessions, it will be passed through more.

The second key element is the free-entry condition, which produces a different response for positive and negative shocks. As all newly formed matches have the same initial idiosyncratic firm productivity, free-entry forces the expected value of posting a vacancy to equate its cost in equilibrium. Consequently, upon creating a new employment relationship, there is a maximum utility level that an entering firm can promise to the worker above which the match would not be viable. Because only new entrants can poach workers, any incumbent offering this value faces zero probability that the worker leaves. This restricts the pass-through but does not apply to negative shocks, thus generating asymmetry. In turn, as the incentives to retain workers for firms exposed to positive shocks are on average high enough to make them update the terms of the employment contract towards a zero poaching probability regardless of the business cycle in the used calibration, this drives down the magnitude of positive changes in promised utility and wages, and it does so in a similar fashion for recessions and non-recessions, which enables to replicate the non-state dependence for these shock.

Besides matching untargeted earnings pass-through asymmetries, the model provides a new explanation for the business cycle trends in income risk documented by a recent empirical literature (Guvenen et al., 2014; Busch et al., 2022) and confirmed also in this paper. Specifically, since the pass-through of negative shocks is on average larger in recessions and that of positive ones is acyclical, the model-generated wage growth distribution for stayers exhibits procyclical skewness and acyclical variance.

Finally, I use the model to evaluate the welfare cost of business cycles. Workers would be willing to give up a significant fraction of their consumption to avoid the effects of aggregate uncertainty in this setting, up to 2.7 percent. While this is a relatively large number compared to the estimates found by other studies (Lucas, 1987; Krusell et al., 2010), and some caution is required in interpreting this result because of non-trivial features in my framework, the lack of a saving technology in the model and the presence of firm disasters make the repercussions of wage fluctuations particularly adverse.

Relation to the literature. This paper relates to several strands of the literature. First, it contributes to the empirical research studying how workers' earnings are affected by idiosyncratic firm shocks. A seminal contribution in this area is from Guiso et al. (2005), who estimate the pass-through of persistent and transitory idiosyncratic firm shocks on stayers using matched employer-employee data from Italy. They find that workers are overall well insured: salaries are almost entirely insulated from transitory shocks and their elasticity with respect to permanent ones is very low, around 0.07. Subsequent papers have replicated their study in other countries, finding remarkably similar estimates (for a review, see Guiso and Pistaferri, 2020).[4] More recently, the increased availability of high-quality administrative data has enabled researchers to overcome some limitations related to the sample selection of stayers (Chan et al., 2020; Friedrich et al., 2019). This literature, however, has in general abstracted from business cycles. An exception is Chan et al. (2020), which is the closest contribution to this paper. Using Danish administrative data on wages, they estimate pass-through regressions to gauge the wages' elasticity to idiosyncratic firm TFP shocks. In contrast

[4]Other papers have also investigated pass-through heterogeneity in worker and firm characteristics (Juhn et al., 2018) and studied differences in the transmission of firm-level and industry-level shocks (Carlsson et al., 2016).

with this paper, they find that persistent negative shocks are always passed through regardless of the state of the cycle, whereas positive ones are only passed through in non-recessions. While data limitations do not allow me to check in full the effects of the diversities between their approach and mine, my paper contributes to this literature by providing results for alternative measures of workers' compensation and institutional context - earnings and Sweden, respectively - and by showing that these differences have non-trivial effects on the results. In addition, rather than reporting only the average pass-through, I document how earnings change over the whole distribution of idiosyncratic firm shocks.

Second, this study relates to the theoretical literature on employment contracts (started by Baily, 1974; Azariadis, 1975) and its application in directed search models to study labor market phenomena (e.g., Menzio and Shi, 2011, 2010; Schaal, 2017; Rudanko, 2009). In particular, it builds on recent work by Balke and Lamadon (2022), who rationalize the pass-through of firms' shocks in a framework in which the main force is firms' trade-off between insuring workers and shaping their incentives to search while on the job. I extend their mechanism to an economy with business cycles and show that adding asymmetries in the stochastic process governing firms' shocks is crucial for matching the patterns in the data.

Third, this paper is connected to studies analyzing, respectively, business cycle heterogeneity in the distributions of firms' productivity shocks and income shocks. Specifically, I confirm in Swedish data the finding by Salgado et al. (2020) that the skewness of firms' productivity shocks is procyclical. I also document, as do Bloom et al. (2018) and Carlsson et al. (2022), the countercyclicality of the standard deviation of these shocks for manufacturing firms. However, while I still find that volatility is slightly higher in recessions, no clear business cycle trend for it is visible in the full sample, which also includes companies operating in retail, construction, and services, as this quantity is high in some non-recessionary years too. Regarding income shocks, I document their procyclical skewness and acyclical variance in my sample, which includes only workers with stable employment relationships. This resonates nicely with the seminal contribution by Guvenen et al. (2014) who report similar trends in US administrative data and, in particular, with Busch et al. (2022) who corroborate these findings for Sweden, Germany and France and for continuously employed full-time workers.[5] In addition, since the model is able to generate these features in the wage growth distribution, my paper contributes to the part of this literature that studies explanatory mechanisms for these patterns. One of the main references in this area is Hubmer (2018), who shows that a job ladder model in a frictional labor market can also replicate the asymmetries found in the data.

Structure of the paper. The paper is structured as follows. Section 1.2 presents the empirical findings, Section 1.3 describes the model, Section 1.4 deals with model calibration, Section 1.5 presents the results obtained from the model and Section 1.6 concludes.

1.2 Empirical analysis

The empirical analysis proceeds in two steps. First, I analyze idiosyncratic firm productivity shocks. I document that negative shocks have larger and more persistent effects on the firm's profitability if they occur in recessions. Second, I examine the relationship between these shocks and workers' earnings. Specifically, I estimate a statistical model of earnings and show that residuals' log changes over the firm's TFP shock distribution exhibit different patterns in non-recessions and downturns.[6]

[5] For Spain, Arellano et al. (2021) find not only that skewness decreases in recessions but also that the variance rises.

[6] In institutional settings in which wage bargaining is highly centralized, firms have less leeway in transferring shocks to workers. Appendix 1.B briefly describes some features of the Swedish labor market context to show that a substantial part of agreements involves firm-level bargaining.

1.2.1 Data

The empirical part of this paper is based on a matched employer-employee dataset created by merging information contained in three different administrative databases assembled by Statistics Sweden and accessed through the servers of the Institute for Evaluation of Labor Market and Education Policy (IFAU), in Swedish *Institutet För Arbetsmarknads-och Utbildningspolitisk Utvärdering*.

The first is the Structural Business Statistics (SBS) dataset, which mainly contains balance sheet and accounting information for the universe of non-financial corporations in Sweden starting from 1997. From this dataset, in addition to balance sheet items, I recover a value-added measure constructed by Statistics Sweden and information on the operating sector of each firm.

The second source is the Register-Based Labour Market Statistics (RAMS) dataset, which contains information on the universe of employment relationships in Sweden from 1985 onward. On the firm's side, I collect information on the type of legal entity and the municipality where the company is located. On the workers' side, I gather information on salaries, length of working relations, and the worker's occupational status (employee, self-employed, etc.). Then, I classify a worker as employed if she is working at least six months in a year, and I assign each employed worker a unique working place each year as the firm where she gets the highest salary. Based on this, I construct an annualized salary measure when a worker stays at the firm for less than a full year and compute firm-level total employment.

The last dataset is the Longitudinal Database on Education, Income and Employment (LOUISE), which contains information on the socio-economic and demographic status of the Swedish population from 1985 onward. From this source, I recover information on civil status, gender, year of birth, number of children, and education.

Sample restrictions. The time period of the analysis is 2004-2018.[7] On the firm side, I restrict the sample to firms with at least five employees, with positive value-added, wage bill, total fixed assets, and equity, whose juridic form is limited liability company, and operating in manufacturing, construction, retail, and services sectors. On the worker side, I consider individuals between 20 and 60 years old whose occupational status is employee. To limit the impact of outliers, I exclude workers in the bottom 5 percent of the distribution of real[8] annual earnings (around 55,000 SEK) and firms in the bottom 0.5 percent of the distributions of value-added (around 515,000 SEK) and wage bill (around 465,000 SEK). Appendix 1.B contains more information on the data.

Sample description. Figure 1.1 shows sample coverage over time. The number of firms included is around fifty thousand each year, and the number of workers is about 1.5 million at the beginning and gradually increases roughly to 1.8 million. As visible from the first two rows, the Great Recession had a significant impact on the Swedish economy between 2009 and 2010.

Looking at different sectors, Figure 1.1 reveals that about one-third of the firms operate in services, followed - in decreasing order - by retail, manufacturing, and construction. A similar pattern holds on the workers' side, with the difference that manufacturing employs more people than retail. In terms of size, more than two-thirds of the firms are small. Medium-size firms (20-49 employees) come second, followed by large (50-99) and very large firms (100 or more). Despite being few, the latter group employs a substantial part of workers - about two-thirds, followed by small, medium, and large companies.

The bottom part of Figure 1.1 considers sample coverage in terms of total value-added and employment in manufacturing, construction,

[7] In 2003, the value-added measure in my data is not reliable for construction and services. The last year in the merged dataset is 2018.

[8] Throughout the paper the reference year for variables in real terms is 2020.

retail, and services, computed from national accounts.[9][10] In all years, the sample covers around 65 percent of value-added and employment. Manufacturing and services constitute the main bulk, followed by retail and construction. When looking at size, the largest share for both value-added and employment is made of very large firms followed, in decreasing order, by small, medium, and large companies. Overall, the sample provides good coverage of the sectors considered.

Tables 1.1 and 1.2 present some summary statistics for workers and firms computed pooling across all the years in the sample. Overall, there are almost 4 million unique workers, 26 million worker-year observations, and about 130 thousand unique firms and 800 thousand firm-year observations.

Considering workers first, Table 1.1 reveals that the average worker in the sample is around forty years old, has a 0.55 probability of having children, and has about two kids, conditional on having any. Roughly one-third of the workers are married. The same proportion has more than a high school diploma. On average, workers stay about six years at the same firm, but there is sizable variability. Males are over-represented, being just less than two-thirds.

The average salary is about 370 thousand SEK, but there is considerable heterogeneity, as the standard deviation is just below the same number. Average earnings are increasing in firm size, except for very large firms in retail and services. Conditional on working in small and medium-sized firms, the average salary is similar across sectors. Workers at large or very large construction firms or at very large manufacturing companies get paid relatively more. Earnings' heterogeneity, in general, also increases in firm size and is highest in services and lowest in construction.

[9] Appendix 1.B describes the sources used for aggregate data.

[10] Figure 1.A.1 provides the same shares as a fraction of totals for the non-financial private sector and the whole economy.

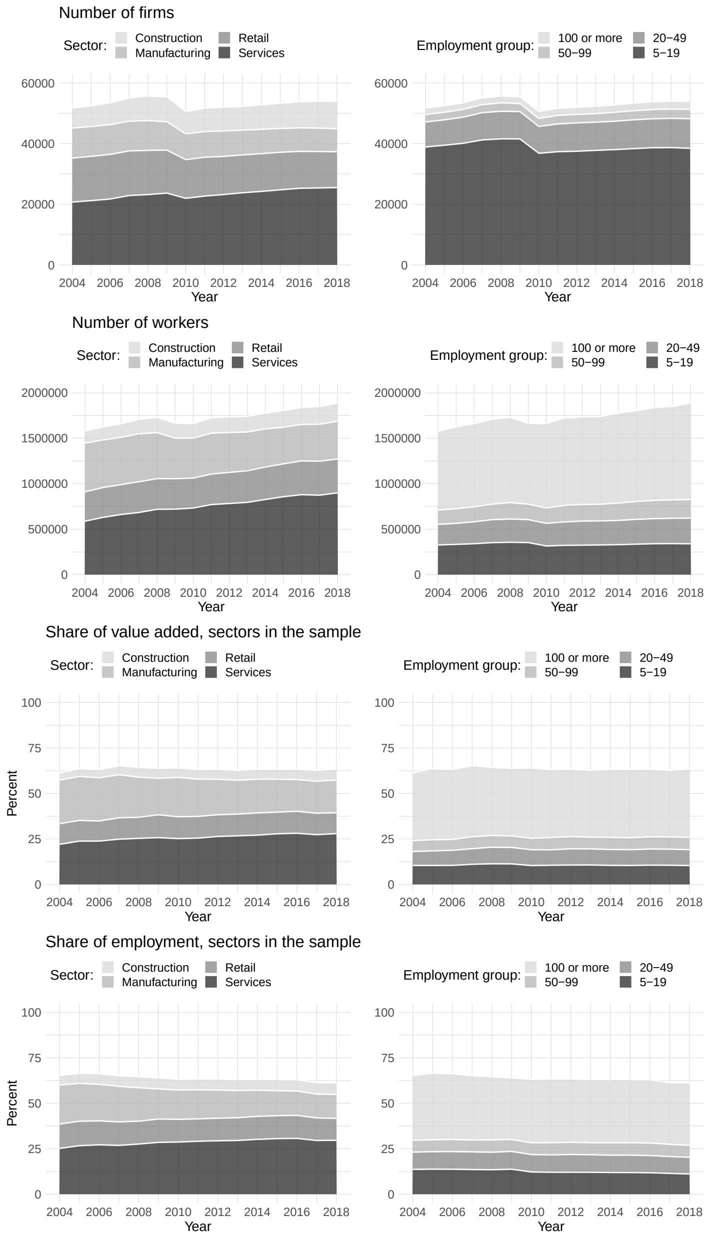


Figure 1.1: Sample coverage by sector and employment group. The third and fourth rows depict, respectively, the share of aggregate value-added and employment in the sectors in the sample (manufacturing, construction, retail, services).

	Employment group				
	5-19	20-49	50-99	100+	All
I. Manufacturing					
Earnings	331,672	356,344	379,223	429,605	404,865
	(211,802)	(222,773)	(262,543)	(337,988)	(309,148)
II. Construction					
Earnings	335,381	374,057	408,851	439,798	388,421
	(128,985)	(149,863)	(231,959)	(224,314)	(188,520)
III. Retail					
Earnings	324,741	370,029	374,333	347,059	348,145
	(213,866)	(292,625)	(303,479)	(315,919)	(288,843)
IV. Services					
Earnings	335,250	363,043	374,706	364,981	360,175
	(292,137)	(304,042)	(303,493)	(332,365)	(318,283)
I. All					
Earnings	332,019	364,701	378,712	387,277	372,315
	(239,902)	(269,209)	(287,816)	(327,419)	(300,525)
Age	38.9	(11.2)			
Share with children	0.55				
Number of children, conditional	1.8	(0.8)			
Share married	0.36				
Share with more than high school	0.33				
Tenure, years	6.1	(5.8)			
Share females	0.36				
Unique workers	3,970,335				
Worker-year observations	25,986,135				

Table 1.1: Summary statistics for workers, 2004-2018. St. dev. in parenthesis. Monetary values in 2020 SEK.

On the firm side, Table 1.2 shows that the value-added per worker for the average firm in the sample is about 760 thousand SEK, and its growth rate is 3.5 percent. However, as depicted by the standard deviations, there is substantial heterogeneity. Value-added per worker is increasing in firm size. Its growth rate is, instead, decreasing overall. The services sector has the highest variability in value-added per worker, and construction has the lowest.

	Employment group				
	5-19	20-49	50-99	100+	All
I. Manufacturing					
VA/worker	731,738	791,437	918,059	1,077,034	787,303
	(842,344)	(1,809,483)	(2,397,940)	(1,937,792)	(1,379,942)
VA/worker, log growth	0.036	0.014	0.009	0.006	0.026
	(0.341)	(0.323)	(0.361)	(0.346)	(0.339)
II. Construction					
VA/worker	705,033	731,824	785,255	808,999	713,242
	(543,921)	(408,260)	(323,921)	(297,303)	(516,778)
VA/worker, log growth	0.052	0.020	0.019	−0.004	0.044
	(0.313)	(0.287)	(0.290)	(0.327)	(0.308)
III. Retail					
VA/worker	707,867	828,028	843,325	885,480	737,479
	(708,197)	(805,108)	(886,850)	(1,219,887)	(756,142)
VA/worker, log growth	0.032	0.006	0.004	0.000	0.025
	(0.348)	(0.347)	(0.365)	(0.341)	(0.349)
IV. Services					
VA/worker	768,086	776,289	815,412	870,022	776,861
	(1,867,334)	(1,256,015)	(1,246,819)	(2,324,583)	(1,773,234)
VA/worker, log growth	0.053	0.014	0.019	0.006	0.040
	(0.370)	(0.345)	(0.342)	(0.320)	(0.361)
V. All					
VA/worker	736,856	784,384	844,613	932,135	759,328
	(1,340,623)	(1,232,919)	(1,530,771)	(1,974,596)	(1,368,924)
VA/worker, log growth	0.045	0.013	0.013	0.004	0.035
	(0.351)	(0.334)	(0.348)	(0.332)	(0.347)
Unique firms	129,500				
Firm-year obs.	798,610				

Table 1.2: Summary statistics for firms, 2004-2018. St. dev. in parenthesis. Monetary values in 2020 SEK.

1.2.2 Recessionary and non-recessionary episodes

Figure 1.2 plots the growth rate of real GDP and the unemployment rate in Sweden.[11] GDP growth was negative just in three years, 2008, 2009, and 2012. Lagged unemployment rose in the aftermath of the Great Recession, decreased afterward, and then slightly rose again during the

[11]See Appendix 1.B for more details on the data.

European debt crisis. Growth slowed down - but was still positive - from 2016 until the end of the sample. During the same period, unemployment gradually decreased.

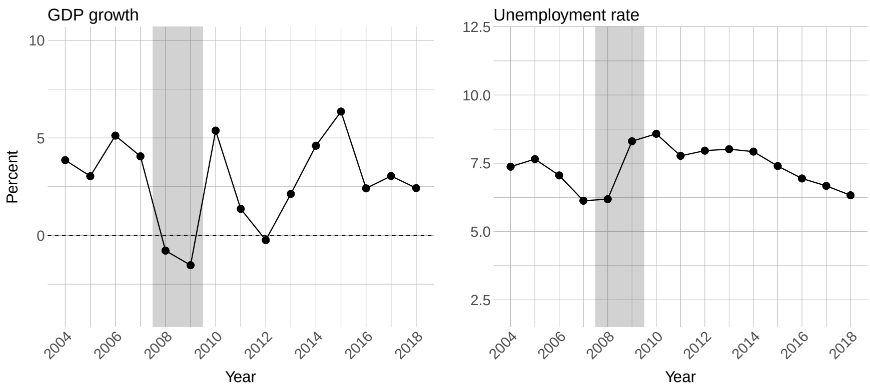


Figure 1.2: GDP growth and unemployment in Sweden. Shaded area corresponds to recessionary years.

Based on these graphs, I define recessions (R) the years 2008-2009 and non-recessions (NR) the remaining years. I do not include 2012 because lagged unemployment did not rise significantly in that episode. As described in Appendix 1.C, I have also experimented with alternative definitions.

1.2.3 Idiosyncratic firm shocks

Following Friedrich et al. (2019), I use the logarithm of value-added per worker as measure of productivity.[12] I denote this variable for firm j at time t with z_{jt}. Then, as in Bloom et al. (2018), I estimate the following model:

$$z_{jt} = \rho z_{jt-1} + \lambda_t + \mu_j + \nu_{jt} \tag{1.1}$$

[12] I have also experimented using the residual from the decomposition of a Cobb-Douglas production function including both labor and capital. As shown in Appendix 1.C, results are robust to this alternative definition.

where λ_t are time fixed effects, μ_j firm fixed effects, and ν_{jt} the residual, which is the measure of idiosyncratic TFP shock.[13]

To better understand the properties of the shocks over the business cycle, Figure 1.3 plots their cross-sectional standard deviation, Kelly's skewness over time, and the difference between the cumulative density function in recessions and non-recessions. Even if the magnitudes are not as large as previously found by the literature (Bloom et al., 2018; Salgado et al., 2020; Carlsson et al., 2022), it is evident in the pictures that the Great Recession was a period of higher uncertainty and of increased risk of being exposed to negative shocks. However, unlike the above-mentioned studies, Figure 1.3 does not support a clear countercyclical business cycle trend in the standard deviation of the shocks in my sample.[14]

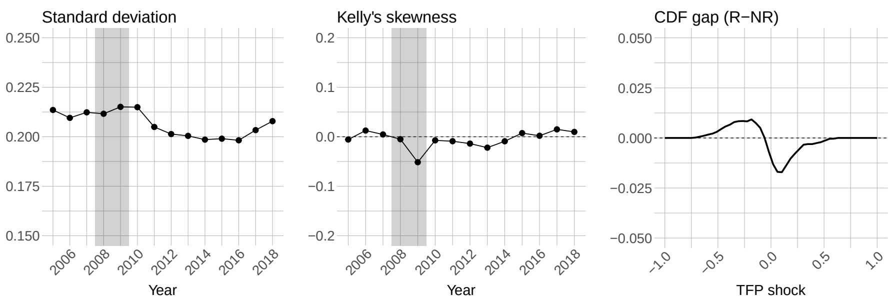


Figure 1.3: Measures of cross-sectional dispersion of idiosyncratic firm TFP shocks over time and CDF difference between recessions and non-recessions.

Figure 1.4 takes a closer look at business cycle asymmetries in the

[13]I acknowledge - like Bloom et al. (2018) - that this measure potentially includes shocks other than pure TFP ones. Two elements, nevertheless, are reassuring that it is a good proxy for them. First, as explained in footnote 12, results are robust when controlling for capital. Second, while I do not have information on prices and capacity utilization at the firm level, a recent paper by Carlsson et al. (2022) focusing on the Swedish manufacturing sector, finds that controlling for these two factors when constructing productivity results in TFP shocks with similar properties to the ones obtained when not doing that. Specifically concerning prices, the authors find a limited impact because they do not change much following pure TFP or demand shocks.

[14]Results available upon request show, instead, that the standard deviation of the shocks is countercyclical when restricting the sample to manufacturing firms.

impact of idiosyncratic shocks on firms' performance by reporting the difference in the share of firms experiencing mass layoffs[15] between recessions and non-recessions over the distribution of TFP shocks.[16]

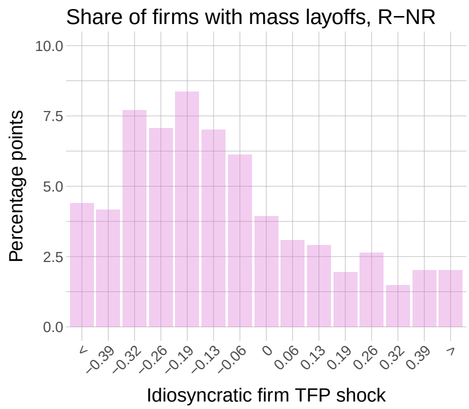


Figure 1.4: Difference in the share of firms experiencing mass layoffs over the distribution of idiosyncratic firm TFP shocks between recessions and non-recessions. The first bin contains all firms below the -0.39 bin and the last all firms above the 0.39 bin.

Compared to those exposed to positive ones, firms experiencing negative shocks in recessions are more likely to face mass layoffs. Under the assumption that mass layoffs are good proxies for long-lasting unfavorable circumstances (it takes on average 2.8 years for a firm experiencing a mass layoff after a negative shock in recession to return to positive employment growth), the graph provides evidence that negative idiosyncratic firm shocks are more disastrous events if they occur in recessions.

[15]Following von Wachter et al. (2009), I define a mass layoff as a reduction of more than 30 percent of the workforce in the two years following the shock.

[16]Figure 1.A.2 in the Appendix reports the shares for recessions and non-recessions in levels.

1.2.4 Workers' earnings

To recover the part of the change in salary not due to observables, I use a statistical model of worker earnings. Let w_{ijt} indicate log earnings of individual i employed at firm j in period t. I assume that w_{ijt} can be modelled with the following linear specification:

$$w_{ijt} = X'_{ijt}\phi + \omega_{ijt} \tag{1.2}$$

The matrix of control variables X_{ijt} includes worker- and firm-specific characteristics and information related to each employment relationship. More in detail, on the worker side, I include a third-order polynomial of age, a dummy variable for males, a dummy variable for having children, dummy variables for four education levels (pre-secondary, high school, post-secondary and post-graduate) interacted with dummies for age groups (20-29, 30-39, 40-49 and 50-60), and dummy variables for four civil statuses (single, married, separated, survivor). On the firm side, I control for sector-specific time trends with sector-year fixed effects[17], for location-specific factors with dummy variables for the region where the firm is based and for firm size with dummy variables for the firm's employment group category (5-19, 20-49, 50-99, 100 or more workers). Finally, I include a third-order polynomial of the tenure of each firm-worker employment relationship.

Following the original approach by Guiso et al. (2005), I focus on workers with steady employment and tenure histories. Therefore, the specification in (1.2) is estimated only on stayers, which I define as workers who remain employed at the same firm for more than two consecutive years. The measure of residual log earnings change I consider is then $\Delta\omega_{ijt} := \omega_{ijt} - \omega_{ijt-1}$.[18]

[17] I use the one-letter sector classification provided by Statistics Sweden. One-letter sectors are broader than two-digit sectors. For instance, the letter M sector belongs to services, includes seven two-digit sectors, and its description is *"Professional, scientific and technical activities"*.

[18] It is worth noting that, while I am interested in the variation due to firm-related shocks, ω_{ijt} contains potentially both firm-related and worker-related shocks. How-

Figure 1.5 plots the standard deviation and Kelly's skewness of $\Delta\omega$ over time. As it is clear from the picture, the standard deviation exhibits no cyclicality while skewness turns negative during the Great Recession, which resonates nicely with the findings in Guvenen et al. (2014) and Busch et al. (2022).

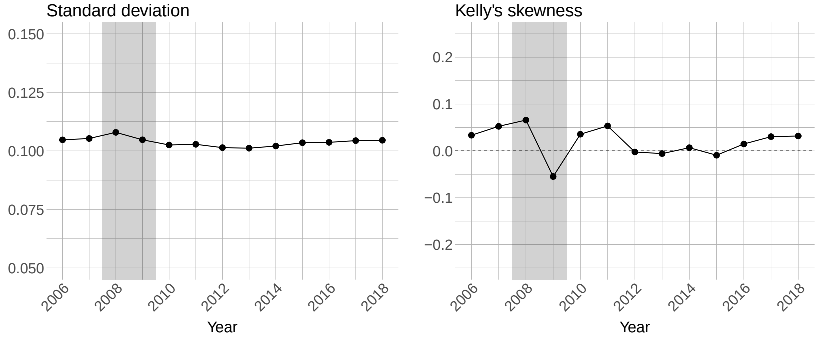


Figure 1.5: Measures of cross-sectional dispersion of residual log earnings changes over time.

1.2.5 Earnings changes and idiosyncratic firm shocks

In this section, I investigate the relation between the measures of idiosyncratic firm shocks (ν) and of residual log earnings changes ($\Delta\omega$).[19] To this end, I classify firms in fifteen bins - defined symmetrically around zero - according to the size of the shocks. Then, I compute the mean residual log earnings change $\Delta\omega$ for all the workers employed at firms in the bin. Figure 1.6 presents the results.

ever, because I am interested in the cross-sectional average of $\Delta\omega$ by firm TFP shock bins, the worker-related part should be at least partially controlled for.

[19]To limit the impact of outliers, idiosyncratic firm shocks are winsorized at 1 and 99 percent and residual log earnings changes at 5 and 95 percent.

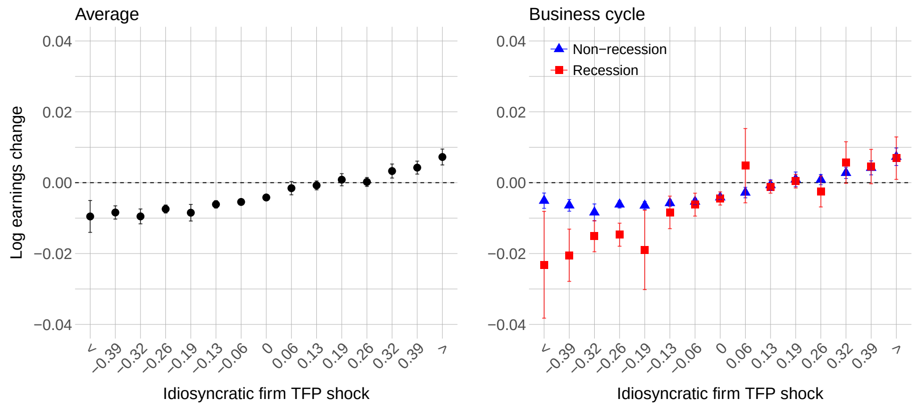


Figure 1.6: Average residual log earnings change over bins of idiosyncratic TFP firm shocks. *Left*: all years. *Right*: recessions (squares) and non-recessions (triangles). Confidence intervals are at the 90 percent level and standard errors are clustered at the firm level. The first bin contains all firms below the -0.39 bin and the last all firms above the 0.39 bin.

Considering first the left part of the figure, which depicts pooled results, it is easy to see that, in general, firms pass through negative shocks. Nevertheless, the gradient is more pronounced for shocks closer to zero and gradually becomes flatter. The opposite pattern holds for positive ones, the pass-through of which is, thus, mainly driven by sizeable shocks. These findings are in line with the positive and significant pass-through coefficients found by the literature (e.g., Guiso et al., 2005; Chan et al., 2020; Guiso and Pistaferri, 2020).

The right panel, however, reveals that pooled results hide substantial business cycle heterogeneity. The pass-through of positive shocks is not state-dependent, as the pattern for recessions and non-recessions replicates the average case. On the other hand, firms provide, in general, significant insurance against negative shocks, but they pass them through much more in downturns.[20]

[20]These findings contrast with the results obtained by Chan et al. (2020) using Danish administrative data on wages: in their paper, negative permanent shocks are always passed through regardless of the state of the cycle, while positive ones are passed through only in non-recessions.

In order to give some magnitude to the trends shown in Figure 1.6, Table 1.3 shows the results obtained from running a regression of the residual log earnings change $\Delta\omega_{ijt}$ on firm shocks ν_{jt} when considering, respectively, all years, recessions and non-recessions.

Dependent variable: $\Delta\omega_{ijt}$, residual log earnings change						
	Average		Recession		Non-recession	
ν_{jt}	0.018^{***}	0.023^{***}	0.035^{***}	0.020^{***}	0.014^{***}	0.024^{***}
	(0.002)	(0.002)	(0.006)	(0.007)	(0.001)	(0.002)
$\nu_{jt} \cdot \mathbb{1}(\nu_{jt} < 0)$		-0.010^{**}		0.026		-0.018^{***}
		(0.004)		(0.020)		(0.003)
Obs. (Millions)	11.4	11.4	1.7	1.7	9.7	9.7
Monetary value (SEK)						
ν_{jt}	1,381	1,764	2,693	1,539	1,073	1,840
$\nu_{jt} \cdot \mathbb{1}(\nu_{jt} < 0)$		997		3,540		460

Table 1.3: The table reports the results from regressing residual log earnings changes on firm shocks. The first column reports the results when the sample includes all years, the second considers recessionary years and the third non-recessionary years. Clustered standard errors at the firm level in parenthesis. Confidence levels: 0.1 (*), 0.05 (**), 0.01 (***). Monetary values in 2020 SEK.

Before describing the results, it is worth noticing that the extreme cases of full and no insurance are represented, respectively, by zero and unitary regressions' coefficients. A number in between, thus, corresponds to different degrees of partial insurance.

The average elasticity of residual log earnings with respect to firms' shocks is 0.018, which is in line with 0.014, the pass-through coefficient of persistent shocks reported in the literature review by Guiso and Pistaferri (2020) for Sweden.[21] This implies a change in workers' average annual earnings of around 0.37 percent for a firm receiving a one standard deviation shock (about 0.20 log points on average). In

[21] Which, in turn, is taken from a previous version of Balke and Lamadon (2022).

monetary terms[22], this corresponds to 1,381 SEK. The limited magnitude of these numbers implies that firms, in general, are good insurance providers for their workers. Allowing the specification to accommodate an asymmetry between positive and negative shocks, reveals that the elasticity of the latter (which is the sum of the coefficient of ν_{jt} and the interaction term) is smaller, which reflects the previously highlighted gradient differences between shocks of different sign.

Concerning business cycle asymmetries, the results for non-recessions are - unsurprisingly, since almost all the years in the sample are classified as such - very similar to the average case. In recessions, instead, while the estimated elasticity of positive shocks is very similar, that of negative ones is much higher.[23] In monetary terms, the estimated coefficient of 0.020 (adopting the conservative approach of using the statistically significant part) corresponds to 1,539 SEK, which amounts to a non-negligible figure at the aggregate level, given that almost 15 percent of firms in the sample (which employ about 15 percent of workers) face a negative shock larger than one standard deviation in the recession.[24]

Robustness and heterogeneity. In Appendix 1.C, I show that the empirical findings are robust to (i) the usage of different definitions of business cycle episodes and to (ii) the usage of an alternative measure of firm productivity that takes into account capital inputs used for production. I also investigate heterogeneity in the results in terms of firms' size, workers' age, and workers' tenure. Overall, the empirical patterns hold in all these dimensions, even though the business cycle asymmetry for negative shocks is slightly more significant for larger firms and

[22]To get the monetary equivalents, I multiply the relevant coefficient by one standard deviation shock and by average earnings in each subsample considered.

[23]Note that, even if the coefficient for negative shocks in recessions is not significantly different from zero at the usual significance levels, the elasticity is anyways more than three times larger in recessions even when using the same estimated coefficient for both positive and negative shocks (0.020 vs. 0.006).

[24]When using 0.046 as elasticity, this amount is 3,540 SEK, which is even less negligible.

more tenured workers.

Summing up. The empirical part of this paper has shown two facts. First, the transmission of negative idiosyncratic firm TFP shocks to workers' earnings is asymmetric over the business cycle: firms pass through these shocks much more in recessions. Positive shocks, on the other hand, are passed through especially if large, and this holds regardless of the state of the economy. Second, negative idiosyncratic shocks are more disastrous events for firms' profitability if they occur in recessions. The model outlined below will help to rationalize these empirical patterns.

1.3 Model

In this section, I develop a directed search model of the labor market that enables the investigation of idiosyncratic firm shocks' transmission to workers' wages. The model is very similar to the original framework developed by Menzio and Shi (2010), and its purpose of understanding the impact of firm shocks on labor market outcomes is close in spirit to the works of Balke and Lamadon (2022) and Schaal (2017).

1.3.1 Environment

Agents and markets. The model economy is inhabited by a continuum of infinitely lived ex-ante identical workers and a positive measure of firms.

Firms maximize the present discounted value of profits at rate $\beta \in (0,1)$ by transforming with a constant returns to scale technology one unit of labor into $f(y,z)$ units of output. Aggregate productivity is denoted by $y \in Y = \{y_1, \ldots, y_{N_y}\}$ with N_y finite and $\underline{y} = y_1 < \cdots < y_{N_y} = \overline{y}$. Idiosyncratic firm productivity is denoted by $z \in Z = \{z_1, \ldots, z_{N_z}\}$ with N_z finite and $\underline{z} = z_1 < \cdots < z_{N_z} = \overline{z}$.

Workers derive utility from consumption according to the function $u(\cdot)$ - with $u : \mathbb{R} \rightarrow \mathbb{R}$ twice continuously differentiable, strictly increasing and weakly concave function - and maximize the present discounted value of utility at rate β. A worker can either be employed or not. In the former case, her consumption is equal to the wage paid by the firm where she works; in the latter, it is equal to the unemployment benefit b. There is no saving technology in this economy.

The aggregate state of nature s follows a Markov chain with a finite number of states governed by the transition probabilities $\Pi_s(\hat{s}|s)$.[25] and determines the value of aggregate productivity $y(s)$. Idiosyncratic productivity z also follows a Markov chain with a finite number of states governed by the transition probabilities $\Pi_z(\hat{z}|z, s)$, which are allowed to depend on the aggregate state.

The aggregate state of this economy ψ, therefore, includes the aggregate state of nature s but also the distribution of workers across unemployment and employment states, respectively $g_u \in [0, 1]$ and $g_e : X \times Z \rightarrow [0, 1]$ with $g_e(V, z)$ being the share of workers employed at a firm with idiosyncratic productivity z under a contract that guarantees to the worker an expected lifetime utility of V.

Workers look for jobs and firms post vacancies in a continuum of submarkets indexed by the expected utility the firm promises to give to the worker upon being hired in that submarket $x \in X = [\underline{x}, \overline{x}]$ with $\underline{x} < u(b)/(1-\beta)$ and $\overline{x} > u(f(\overline{y}, \overline{z}))/(1-\beta)$. Market tightness - the ratio of vacancies to workers searching for a job - in each submarket depends on the aggregate state and is denoted by $\theta(\psi, x)$.

There are four stages within each period: separation, search, matching, and production. In the separation stage, existing matches are exogenously destroyed with probability $\delta \in (0, 1)$.[26] During the second stage, firms choose how many vacancies and in which submarket to post them at per-unit cost $k > 0$, and workers - both employed and

[25] The hat denotes variables in the next period.

[26] Differently to Menzio and Shi (2010), my model does not allow for endogenous separations.

unemployed - choose in which submarket to look for a job. Currently unemployed and employed workers search, respectively, with intensities $\lambda_u, \lambda_e \in (0, 1]$. Workers and firms in a given submarket are matched together into an employment relationship according to a constant returns to scale function during the third stage. More in detail, the job finding probability in a submarket with tightness θ is denoted by $p(\theta)$. As standard in the literature, $p : \mathbb{R}_+ \to [0, 1]$ is twice continuously differentiable, strictly increasing, strictly concave and it satisfies $p(0) = 0$ and $p'(0) < \infty$. Similarly, the job filling probability is $q(\theta)$ with $q : \mathbb{R}_+ \to [0, 1]$ twice continuously differentiable, strictly decreasing, convex function with $q(\theta) = p(\theta)/\theta$, $q(0) = 1$, $q'(0) < 0$ and $p(q^{-1}(\cdot))$ concave.[27] All new matches start with the same specific value of idiosyncratic productivity $z_0 \in Z$. In the production stage, each employed worker receives and consumes the wage w specified by the employment contract (described more in detail below) for producing $f(y, z)$ units of output. Unemployed workers consume the unemployment benefit $b \in (0, f(\overline{y}, \overline{z}))$. Finally, before the start of a new period, the new aggregate state of nature $\hat{s}$ and the new idiosyncratic productivity state $\hat{z}$ are drawn. Figure 1.7 depicts graphically the timing of actions in the model.

Contracting. Upon matching, the worker-firm relationship is disciplined by a contract that specifies the full series of wages $\{w_{t+j}\}_{j=0}^{\infty}$ for each possible history of the world (s^{t+j}, z^{t+j}). More specifically, in each period t, the firm chooses the wage path in each possible future history to maximize profits. Following the literature on recursive contracts (see Menzio and Shi, 2010; Balke and Lamadon, 2022, and references cited therein), it is possible to specify the firm problem recursively with the addition of future promised utility $\hat{V}$ as state variable. In other words, in each period, the firm chooses the current wage w and the future expected utility to give to the worker in each future state, that is $\hat{V}(\hat{\psi}, \hat{z})$. Because - as explained below - workers' optimal searching choice while

[27] As explained in Menzio and Shi (2010), this assumption is needed to guarantee that the worker's problem is strictly concave and, hence, that it has a unique solution.

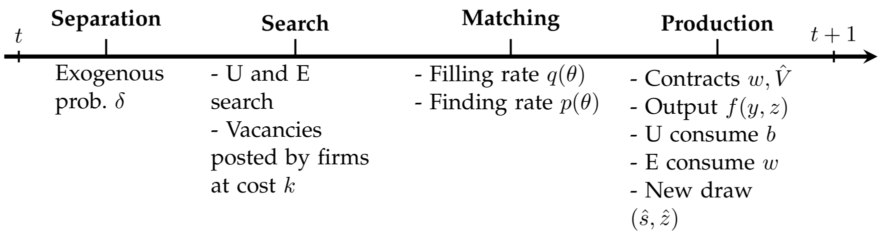


Figure 1.7: Timing of actions.

on the job is a function of the future utility they would get in the current employment relationship, $\hat{V}(\hat{\psi}, \hat{z})$ will also be the relevant variable influencing their on the job choice upon realization of that future state. Commitment is only on the firm side: firms have to give the worker the expected utility specified by the contract - but can choose the split between wages today and future expected utility tomorrow. Workers, instead, are free to choose where to search while on the job. That is, firms will have to choose $\hat{V}(\hat{\psi}, \hat{z})$ so that workers' on-the-job search choices are consistent with the firm's profit maximization.[28]

1.3.2 Worker's problem

Consider first the problem of an employed worker with current promised expected lifetime utility V who has to decide in which submarket x to direct her search while on the job. If she does not get to search, she will get V as specified by the contract. If she searches, instead, with probability $p\left(\theta(\psi, x)\right)$ she finds a new job that will give her expected lifetime utility x while with complementary probability she remains at her current job and gets the future utility specified by the contract. Mathematically, at the beginning of the search stage, the employed worker lifetime utility is $V + \max\{0, R(\psi, V)\}$ where R is

[28]This contractual environment is labeled "dynamic contracts" in Menzio and Shi (2010).

the return to search function defined as:

$$R(\psi, V) = \max_{x \in X} p\left(\theta(\psi, x)\right)(x - V) \tag{1.3}$$

The solution of this problem returns the optimal searching choice $x^*(\psi, V)$ and the probability of leaving $\tilde{p}(\psi, V) := p\left(\theta(\psi, x^*(\psi, V))\right)$.

Let us now turn to the problem of an unemployed worker. At the beginning of the production stage, she consumes the unemployment benefit b and decides where to search in the next period, upon having the possibility to search. Her value is, therefore:

$$U(\psi) = u(b) + \beta \mathbb{E}\left[U(\hat{\psi}) + \lambda_u \max\{0, R(\hat{\psi}, U(\hat{\psi}))\}\right] \tag{1.4}$$

1.3.3 Firm's problem

Consider a firm who is matched with a worker whose lifetime utility specified by the contract is V when the aggregate state of nature is s and the idiosyncratic z. As previously described, conditional on survival of the employment relationship, the firm's problem is to maximize profits by choosing the wage to be paid today to the worker w and the future promised utility in each state of the world $\hat{V}$ subject to delivering the utility promised at the beginning of the period V.

Letting $\xi = (\psi, z)$, then the firm value is given by the following expression:

$$J(\xi, V) = \max_{w, \{\hat{V}(\hat{\xi})\}_{\forall \hat{\xi}}} f(y, z) - w + \beta(1-\delta)\mathbb{E}\left[\left(1 - \lambda_e \tilde{p}\left(\hat{\psi}, \hat{V}(\hat{\xi})\right)\right) J(\hat{\xi}, \hat{V}(\hat{\xi}))\right] \tag{1.5}$$

subject to:

$$V = u(w) + \beta \mathbb{E}\left[\delta U(\hat{\psi}) + (1-\delta)\left(\hat{V}(\hat{\xi}) + \lambda_e R(\hat{\psi}, \hat{V}(\hat{\xi}))\right)\right] \tag{1.6}$$

1.3.4 Market tightness

The firm chooses during the search stage how many vacancies to create and in which submarket to locate them. Its optimal choice on how many vacancies to create in submarket x when the aggregate state is ψ is determined by the benefit of creating a vacancy $q(\theta(\psi, x))J(\psi, z_0, x)$ and the cost of doing that k. The optimal strategy is therefore to create infinite vacancies in submarkets where the benefit is strictly higher than the cost and zero vacancies in submarkets where the cost is strictly higher than the benefit. Upon equality of the benefit and the cost, the firm is indifferent. Therefore, in any submarket that is visited by a positive amount of workers the tightness function is consistent with firm's optimal strategy if and only if:

$$k \geq q(\theta(\psi, x))J(\psi, z_0, x) \tag{1.7}$$

and $\theta(\psi, x) \geq 0$ with complementary slackness. In markets not visited by any worker, consistency requires $q(\theta(\psi, x))J(\psi, z_0, x)$ to be smaller or equal than k. However, following Menzio and Shi (2010), attention is restricted to equilibria in which the following slackness condition is satisfied in every submarket.

1.3.5 Equilibrium

The economy described above admits the following definition of equilibrium:

Definition 1.3.1. *A recursive equilibrium is a market tightness function* $\theta : \Psi \times X \to \mathbb{R}_+$, *a return to search function* $R : \Psi \times X \to \mathbb{R}$, *a search policy function* $x^* : \Psi \times X \to X$, *a value function for unemployment* $U : \Psi \to \mathbb{R}$, *a firm value function* $J : \Psi \times Z \times X \to \mathbb{R}$, *a wage policy function* $w : \Psi \times Z \times X \to \mathbb{R}$, *a future promised expected utility function* $\hat{V} : \Psi \times Z \times X \to X$ *and a law of motion for the aggregate state* $\Phi_\psi : \Psi \to \Psi$ *such that:*

- *Market tightness* θ *is consistent with the firm optimal creation strategy* (1.7) *for all* $(\psi, x) \in \Psi \times X$;

- *Return to search R solves the problem in* (1.3) *and x^* is the associated policy function for all $(\psi, V) \in \Psi \times X$;*
- *Unemployment value satisfies* (1.4) *for all $\psi \in \Psi$;*
- *Firm value function solves the problem* (1.5) *and w and $\hat{V}$ are the associated policy functions for all $(\psi, z, V) \in \Psi \times Z \times X$;*
- *Aggregate law of motion of the economy Φ_ψ is derived from the policy functions $(x^*, w, \hat{V})$ and the exogenous processes governing s and z.*

The above definition makes clear that the equilibrium objects of this economy are, in principle, functions also of the distribution of workers across the unemployment/employment states g_u, g_e, which would require solution procedures similar to the one developed by Krusell and Smith (1998) and make the computation of the equilibrium demanding. When the equilibrium depends on the aggregate state just through the aggregate state of nature s and not on the distribution of workers g_u, g_e, we say that such equilibrium is block recursive:

Definition 1.3.2. *A block recursive equilibrium is a recursive equilibrium such that the equilibrium objects $\{\theta, R, x^*, U, J, w, \hat{V}\}$ depend on the aggregate state ψ just through the aggregate state of nature s and not through the distribution of workers g_u, g_e.*

As in Menzio and Shi (2010), directed search implies the following proposition:

Proposition 1.3.3. *If the economy admits an equilibrium, then it is block recursive.*

Proof. See Appendix 1.D. □

Therefore, the equilibrium depends on the aggregate state ψ just through the aggregate state of nature s and not through the distribution of workers g_u, g_e, which considerably simplifies the solution of the model.[29]

[29] Appendix 1.E describes the numerical solution procedure.

1.3.6 Trade-offs

To give some intuition on the main mechanisms in the model, in this section I discuss the trade-offs related to the optimization problems of the worker and the firm, respectively. Considering the worker first, the optimality condition of the maximization problem reads:

$$\frac{\partial p\left(\theta(s,x)\right)}{\partial x}(x-V) = -p\left(\theta(s,x)\right) \tag{1.8}$$

The above expression makes clear the trade-off experienced by the worker: on the one hand she would like to search in submarkets where he would receive higher lifetime utility, on the other, in equilibrium the probability of finding a job is decreasing in lifetime utility.

Turning to the firm, combining the first order conditions of its problem returns for each $\hat{\xi}$:

$$\frac{1}{u'(\hat{w})} - \frac{1}{u'(w)} = -\frac{\partial \log \tilde{p}(\hat{s},\hat{V}(\hat{\xi}))}{\partial \hat{V}} \frac{\lambda_e \tilde{p}(\hat{s},\hat{V}(\hat{\xi}))}{1-\lambda_e \tilde{p}(\hat{s},\hat{V}(\hat{\xi}))} J(\hat{\xi},\hat{V}(\hat{\xi})) \tag{1.9}$$

Equation (1.9) clarifies the trade-off faced by the firm when deciding how much to change wages between two consecutive periods. On the one hand - the left side - changing the wage is costly because the firm would like to insure the worker against wage fluctuations. Indeed, because the worker is risk averse, by guaranteeing a stable wage the firm does not need to pay any risk premium. On the other hand - the right side - the firm benefits from changing the wage tomorrow (through its choice today of the state contingent promised utility tomorrow) because it keeps the worker's on-the-job search incentives aligned with the expected value of the match. Indeed, a firm experiencing a large positive idiosyncratic productivity shock would like to choose a high value for $\hat{V}$ in that state of the world to ensure that the worker is not poached and be able to enjoy the higher profits following from the shock.

Equation (1.9) is also important because it tells us the sign of wage changes. Specifically, since the first derivative on the right side is negative and the second term is positive, the change will either be positive or negative depending on the sign of J tomorrow.

In addition, it is worth noting that the crucial element to generating different relative marginal costs across states are searching frictions: without them, the model would prescribe full wage insurance against idiosyncratic shocks.

1.4 Calibration and estimation

The model is calibrated at the quarterly frequency. Below I describe (i) my choices for the functional forms of utility, production and matching functions (ii) the modeling and calibration strategy for the exogenous stochastic processes (iii) the parameters taken from the literature.

Functional forms. I adopt the following functional forms. For the utility function u I use the standard CRRA form. For the matching functions p and q I follow Menzio and Shi (2010). For the production function f I use an exponential function of aggregate and firm-specific productivities normalized by a constant a. Specifically:

$$u(x) = \frac{x^{1-\sigma} - 1}{1-\sigma}, \quad f(y,z) = ae^{y+z},$$
$$p(\theta) = \theta\,(1+\theta^{\gamma})^{-\frac{1}{\gamma}}, \quad q(\theta) = p(\theta)/\theta = (1+\theta^{\gamma})^{-\frac{1}{\gamma}}$$

Exogenous processes. Following Krueger et al. (2016), the aggregate state of nature s is modelled as a two-state (recession R and non-recession NR) first-order Markov process with transition matrix:

$$\pi_s(\hat{s}|s) = \begin{bmatrix} \pi_{\mathrm{R}} & 1-\pi_{\mathrm{R}} \\ 1-\pi_{\mathrm{NR}} & \pi_{\mathrm{NR}} \end{bmatrix}$$

The transition probabilities are calibrated as follows. Recall that the above process implies that the stationary probabilities of the two states

$\pi_{\mathrm{R}}^{\infty}$ and $\pi_{\mathrm{NR}}^{\infty}$, and the expected length of a recession $\mathbb{E}(\mathrm{R})$ are determined by the following formulas:

$$\pi_{\mathrm{R}}^{\infty} = \frac{1-\pi_{\mathrm{NR}}}{2-\pi_{\mathrm{NR}}-\pi_{\mathrm{R}}}, \quad \pi_{\mathrm{NR}}^{\infty} = \frac{1-\pi_{\mathrm{R}}}{2-\pi_{\mathrm{NR}}-\pi_{\mathrm{R}}}$$

$$\mathbb{E}(\mathrm{R}) = 1\cdot(1-\pi_{\mathrm{R}}) + 2\cdot\pi_{\mathrm{R}}(1-\pi_{\mathrm{R}}) + \cdots = \frac{1}{1-\pi_{\mathrm{R}}}$$

Thus, π_{R} can be calibrated to match the expected length of a recession and, given π_{R}, π_{NR} can be calibrated to match the fraction of time that the economy spends in the recessionary state. In my baseline empirical analysis I consider just the financial crisis - corresponding to the years 2008-2009 - as recession. Therefore, the frequency of recessions in Sweden over the time period 2004-2018, $\pi_{\mathrm{R}}^{\infty}$, is 13.3% and the average length of a recession, $\mathbb{E}(\mathrm{R})$, is 8 quarters. Using the formulas above, this delivers $\pi_{\mathrm{R}} = 0.875$ and $\pi_{\mathrm{NR}} = 0.981$.

The two values of log aggregate productivity y_{NR} and y_{R} with $y_{\mathrm{NR}} > y_{\mathrm{R}}$ are calibrated as follows. Normalizing the unconditional mean to one, the following relation holds:

$$e^{y_{\mathrm{R}}}\pi_{\mathrm{R}}^{\infty} + e^{y_{\mathrm{NR}}}\pi_{\mathrm{NR}}^{\infty} = 1$$

Thus, having a value for the ratio $e^{y_{\mathrm{R}}}/e^{y_{\mathrm{NR}}}$, it is possible to use the above equation to get the two values for aggregate productivity. In the data, the ratio at the end of 2009 is equal to 0.97, corresponding to a drop of 3 percent in real GDP per capita in recession rather than in normal times.[30] This implies that the two values for log aggregate productivity are $y_{\mathrm{NR}} = \log(1.004)$ and $y_{\mathrm{R}} = \log(0.974)$.

Two components discipline the behavior of the log of idiosyncratic firm productivity z. The first is labelled $\tilde{z}$ and follows an AR(1) process:

$$\tilde{z}_{jt} = \rho\tilde{z}_{jt-1} + \epsilon_{jt}, \quad \epsilon_{jt} \overset{\text{i.i.d.}}{\sim} \mathcal{N}\left(0,\sigma^2\right)$$

[30] Compared to what a linear trend estimated on the full time range 2004-2018 would predict.

The second is labelled ζ and follows a two-state Markov chain (with states $\{\bar{\zeta}, 0\}$ - I will refer to the first as "Low" and the second as "High" state - and $\bar{\zeta} \leq 0$) whose transition matrix depends on both the state of the economy and the sign of ϵ_{jt}. Therefore, the transition matrices for ζ are fully determined by eight parameters: the probability of remaining in state L conditional on being in it for non-positive and positive shocks and recessions and non-recessions, respectively $\pi^{-}_{\zeta,\mathrm{R,L}}, \pi^{+}_{\zeta,\mathrm{R,L}}, \pi^{-}_{\zeta,\mathrm{NR,L}}, \pi^{+}_{\zeta,\mathrm{NR,L}}$, and the probability of remaining in state H conditional on being in it for non-positive and positive shocks and recessions and non-recessions, respectively $\pi^{-}_{\zeta,\mathrm{R,H}}, \pi^{+}_{\zeta,\mathrm{R,H}}, \pi^{-}_{\zeta,\mathrm{NR,H}}, \pi^{+}_{\zeta,\mathrm{NR,H}}$.

Firm log productivity z is then equal to $\tilde{z}$ when the second component is in the high state and just equal to $\bar{\zeta}$ otherwise. Mathematically:

$$z_{jt} = \begin{cases} \zeta_{jt} & \text{if } \zeta_{jt} = \bar{\zeta} \\ \tilde{z}_{jt} & \text{if } \zeta_{jt} = 0 \end{cases}$$

This modeling choice is motivated by the fact that the two components serve for two different purposes: $\tilde{z}$ is supposed to capture the average behavior of firm productivity, while ζ should catch the empirical fact that, on average, negative shocks are more disastrous events in downturns.

For this reason, I estimate the parameters governing them separately. Specifically, I estimate ρ and σ with SMM assuming ζ is inactive and matching the average standard deviation of firm productivity z and the average standard deviation (i.e., these are the averages computed using all years without conditioning on the state of the cycle) of the firm shocks ν in the data which are, respectively, 0.527 and 0.206.[31]

On the other hand, the process of ζ requires to calibrate nine parameters: $\bar{\zeta}$, $\pi^{-}_{\zeta,\mathrm{R,L}}$, $\pi^{+}_{\zeta,\mathrm{R,L}}$, $\pi^{-}_{\zeta,\mathrm{NR,L}}$, $\pi^{+}_{\zeta,\mathrm{NR,L}}$, $\pi^{-}_{\zeta,\mathrm{R,H}}$, $\pi^{+}_{\zeta,\mathrm{R,H}}$, $\pi^{-}_{\zeta,\mathrm{NR,H}}$, $\pi^{+}_{\zeta,\mathrm{NR,H}}$. I achieve this by matching moments related to the evidence on mass

[31] Because the model is quarterly and the moments from the data are annual, I aggregate both $\tilde{z}$ and ζ at the annual frequency by taking their value in the last quarter. In addition, I define a year as recessionary if there are at least two consecutive quarters in which the economy is in downturn.

layoffs described in Section 1.2.3. More in detail, I proceed as follows. Because in the data there is not much difference in the duration of being in state L for positive and negative shocks, I set $\pi^-_{\zeta,\mathrm{R,L}} = \pi^+_{\zeta,\mathrm{R,L}} = \pi_{\zeta,\mathrm{R,L}}$ and $\pi^-_{\zeta,\mathrm{NR,L}} = \pi^+_{\zeta,\mathrm{NR,L}} = \pi_{\zeta,\mathrm{NR,L}}$. In addition, I set $\pi^-_{\zeta,\mathrm{NR,H}} = \pi^+_{\zeta,\mathrm{NR,H}} = 1$, which implies that firms can enter in the disaster state only upon receiving negative shocks in recessions, which is the relevant asymmetry in the data to be captured. This leaves five parameters to be estimated: $\bar{\zeta}, \pi^-_{\zeta,\mathrm{R,L}}, \pi^-_{\zeta,\mathrm{NR,L}}, \pi^-_{\zeta,\mathrm{R,H}}, \pi^+_{\zeta,\mathrm{R,H}}$. To get them, I adopt again SMM on the full process - using ρ and σ previously estimated - and match five moments: (i) the difference in the average share of firms experiencing mass layoffs upon receiving negative shocks between recessions and non-recessions (0.070), (ii) the same quantity for firms experiencing positive shocks (0.027), (iii) the average duration[32] of the L state for firms entering in it in recession (2.779 years), (iv) the ratio between the share of firms exiting the L state in three years in recessions and non-recessions (0.888) and (v) Kelly's skewness of ν in 2009 (-0.051).

The lower part of Table 1.4 reports the calibration outcome for the parameters governing the exogenous stochastic processes in the model and Appendix 1.F describes in detail the estimation procedure.

Model parameters. The discount factor β is set as in Menzio and Shi (2010) equal to 0.987 in order to match an annual interest rate of about 5%. The unemployment benefit b is set as in Hagedorn and Manovskii (2008) equal to 0.955. As common in the literature, I normalize the search efficiency of the unemployed to one (e.g., Menzio and Shi, 2010). The separation rate δ is set to 0.022 in order to match the quarterly employment-to-unemployment rate in Sweden reported by Balke and Lamadon (2022). I set the vacancy cost to 0.049, the value used by Hagedorn and Manovskii (2008) divided by twelve as they have a weekly model. The matching function parameter is then set to 0.31 in order to match the unemployment-to-employment rate in Sweden as reported

[32]In the data, I measure exit from L state as the first year - after the first year the firm experiences a mass-layoff - of positive employment growth.

by Balke and Lamadon (2022). Finally, I set λ_e equal to 0.45 in order to match the job-to-job transition rate in Sweden from Balke and Lamadon (2022). The upper part of Table 1.4 summarizes the choices for these parameters.

Model parameters and functional forms

Parameter	Value	Description	Source/Target
Externally calibrated			
σ	1.5	CRRA utility parameter	Balke and Lamadon (2022)
β	0.987	Discount factor	Menzio and Shi (2010)
k	0.049	Vacancy cost	Hagedorn and Manovskii (2008)
b	0.955	Value of non-market activity	Hagedorn and Manovskii (2008)
Internally calibrated			
δ	0.022	Separation rate	E2U rate. *Data:* 0.022. *Model:* 0.022.
λ_e	0.45	Search efficiency, employed	J2J rate. *Data:* 0.026. *Model:* 0.029.
γ	0.31	Matching function parameter	U2E rate. *Data:* 0.170. *Model:* 0.163.
Normalizations			
e^{z_0}	1	Productivity new matches	Standard
λ_u	1	Search efficiency, unemployed	Standard
a	1/1.14	Firm productivity, constant	Average firm prod. in steady state equal to 1

Stochastic processes

Parameter	Value	Description	Target	Data	Model
π_{R}	0.875	Prob. stay in R state	Length recession (quarters)	8	
π_{NR}	0.981	Prob. stay in NR state	Recession frequency	0.133	
$e^{y_{\mathrm{R}}}$	0.974	Aggregate productivity in R	GDP per capita R/NR	0.970	
$e^{y_{\mathrm{NR}}}$	1.004	Aggregate productivity in NR			
ρ	0.979	Autocorrelation firm productivity	Average SD firm productivity	0.527	0.527
σ	0.106	St. dev. firm productivity shocks	Average SD firm shocks	0.206	0.206
$e^{\bar{\zeta}}$	0.620	Disaster state	Kelly's skewness ν in 2009	-0.051	-0.041
$\pi^{-}_{\zeta,\mathrm{R,H}}$	0.879	Prob. stay in H state, R, non-pos. shock	R-NR share firms mass layoffs upon neg. shocks	0.070	0.113
$\pi^{+}_{\zeta,\mathrm{R,H}}$	0.986	Prob. stay in H state, R, pos. shock	R-NR share firms mass layoffs upon pos. shocks	0.027	0.018
$\pi_{\zeta,\mathrm{R,L}}$	0.590	Prob. stay in L state, R	Duration (years) disaster state from R	2.779	1.710
$\pi_{\zeta,\mathrm{NR,L}}$	0.885	Prob. stay in L state, NR	R/NR share firms exiting disaster state in 3 years	0.888	1.041

Table 1.4: Parameters.

1.5 Results

1.5.1 Inspecting the mechanism

To provide some intuition for the model mechanisms that generate the results described in the following sections, Figure 1.8 graphically represents the main elements of the theoretical framework.

Considering first the left part, the graph depicts the relationship between a risk-averse employed worker and a risk-neutral firm. As previously described, the firm is committed to delivering the worker a predetermined amount of promised utility V by choosing how much to give today through the wage w and how much through state-contingent promised utilities tomorrow $\hat{V}$. The optimal contract in this simple setting with no on-the-job search frictions and no business cycles prescribes constant wages. Indeed, by providing the worker full insurance against idiosyncratic shocks, the firm minimizes labor costs as it avoids paying the risk premium for wage fluctuations implied by the curvature of the worker's utility function.

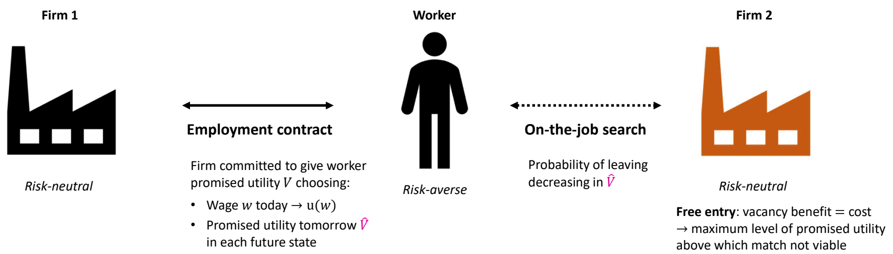


Figure 1.8: Graphical representation of the main elements of the model.

Turning to the model with on-the-job search, the firm now needs to consider that the choice of future promised utilities impacts the worker's probability of leaving. To make a concrete example, consider the case of a firm exposed to a negative idiosyncratic productivity shock. Because the shock substantially reduces the value of the match, the firm has a lower incentive to keep the worker. Thus, it will optimally choose a very low promised utility in that state of the world so that the worker will have a higher probability of being poached.[33]

Figure 1.9 reinforces the point by plotting the model-implied optimal policies of promised utility, probability of leaving and the implied wage equivalent[34] for a firm experiencing an idiosyncratic nega-

[33]Therefore, the mechanism will be stronger the larger and more persistent the shock is.

[34]The wage equivalent of utility V is defined as the constant wage that the firm would need to give to the worker to achieve that utility, that is, is the wage w that

tive shock as functions of their current values (dashed line). When the shock realizes, the firm provides the worker a lower promised utility, which translates into a higher probability of leaving. In addition, because a lower wage profile can sustain the decreased value of promised utility, the firm also cuts wages.

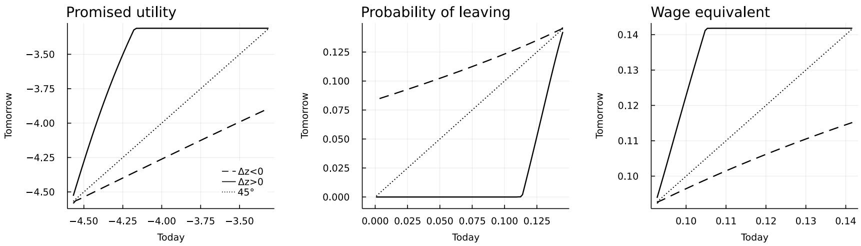


Figure 1.9: Policy functions of promised utility, probability of leaving and wage equivalent as function of their current values.

The graph also depicts the case of a firm exposed to a positive idiosyncratic productivity shock (solid line). While the mechanism works in the opposite direction, it is not symmetric because of the free entry condition. As illustrated in Figure 1.8 and in equation (1.7), the cost of posting a vacancy should equal the benefit of doing that. Hence, a maximum level of utility above which the match for an entering firm would not be profitable must exist. In turn, because only entering firms can poach workers, incumbents offering promised utility equal to the threshold effectively neutralize the effects of on-the-job search and, as prescribed by the frictionless model optimally provide full insurance - see equation (1.9). The implications of this discussion are evident in the discontinuities depicted in the picture. Losing the worker would be so damaging for a firm experiencing a positive shock that it offers the upper bound value even for relatively low values of the current promised utility.

Two facts support this mechanism as a good representation of what happens in reality. First, an immediate testable implication is

---

solves $V = \frac{u(w)}{1-\beta}$.

that the probability of leaving should be a decreasing function of the size of the idiosyncratic firm shock. The right panel of Figure 1.11 shows that the data confirm this prediction. Second, compared to other OECD countries, Sweden features stricter than average employment protection (OECD, 2021). In such a setting, it is not hard to imagine that firms might be more prone to use other mechanisms - for example, incentive-based ones like those analyzed in this paper - rather than dismissal to achieve their employment objectives.[35]

Finally, two are the key factors allowing the model to replicate the business cycle asymmetry found in the data for negative shocks. The first is the general equilibrium effect from the lower job finding probability in recessions. Indeed, achieving the same probability that the worker leaves with a lower job finding rate requires a more significant decrease in promised utility.[36] The second is modeling the increased probability of firm disaster in recessions, documented empirically in Section 1.2.3. Intuitively, this feature reduces even more the value of the match in recessions conditional on the firm being exposed to a negative idiosyncratic productivity shock, which strengthens the mechanism previously explained. The fact that the upper bound of utility binds in both recessions and non-recessions is instead crucial to deliver the non-state dependence of the pass-through of positive shocks.

1.5.2 Steady state

Figure 1.10 plots the average log wage change over the distribution of idiosyncratic firm productivity shocks resulting from simulating the model in steady state. Compared to its corresponding data version, the

[35]The evidence presented in Appendix 1.C that the business cycle asymmetry for negative shocks is stronger for more tenured workers is consistent with this argument given that, in general, these employees are more difficult to fire.

[36]Similarly, achieving the same probability that the worker stays requires a less pronounced increase in promised utility. However, as described below, unlike the scenario in which the firm wants the worker to leave, the general equilibrium effect, in this case, is not very significant in the current calibration.

match is quite good.[37]

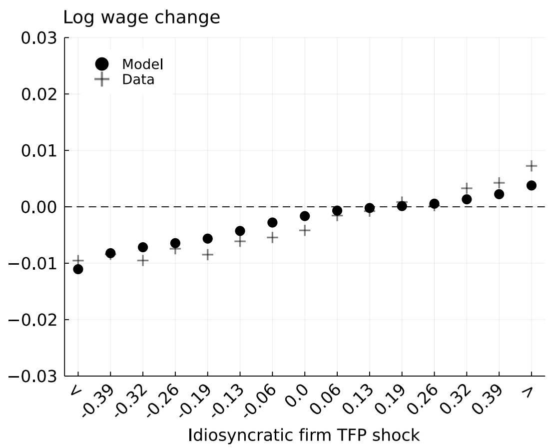


Figure 1.10: Average log wage change in steady state over the distribution of idiosyncratic firm productivity shocks.

As explained in the previous section, when a firm is exposed to a shock, it reoptimizes the worker's contract to align the promised utility with the new expected value of the match. Since the estimated persistence of the shocks is high, large negative shocks substantially reduce the expected value of the match. Thus, upon experiencing one, the firm has a lower incentive to sustain the job relationship, and it optimally increases the chance that the worker is poached by promising a lower utility. In turn, a lower promised utility implies a lower wage. The closer a negative shock is to zero, the weaker the channel just described and, therefore, the smaller the pass-through.[38] Figure 1.11 validates this rea-

[37]In the model, quarterly wages are aggregated at the annual frequency by summing their values over the four quarters, an individual is classified as a stayer upon remaining at the firm for all the four quarters, and a year is defined as recessionary if there are at least two consecutive quarters in which the economy is in downturn. To limit the impact of outliers, the wage growth distribution is winsorized at the top and bottom 1 percent (the same restriction applies for the results in all the sections from here onwards).

[38]Note that, unlike for positive shocks, there is no lower bound for the pass-through of negative ones. Thus, while providing a good match, the model misses capturing the flattening of the slope for large negative shocks in the average case (and, as depicted

soning by showing that promised utility growth is decreasing and the probability of leaving increasing in the size - in absolute value - of the shock.

Regarding positive shocks, Figure 1.10 shows that the model matches well the corresponding patterns in the data. The reason for this is related to the upper bound level of promised utility. For values above that threshold, the poaching probability is zero, because it is not viable for a new firm to start a new match. Thus, since it is profitable for firms receiving positive shocks to keep the relationship alive, they immediately offer the worker values close to the threshold to minimize the probability of leaving. Remembering that the first order condition (1.9) implies that search frictions are the only reason for the lack of full insurance against idiosyncratic shocks, as the firm reduces the poaching probability, it also guarantees a smoother path of wages. Again, Figure 1.11 corroborates this reasoning by showing that promised utility growth increases and the probability of leaving decreases in the size of the shock.

The same picture also well depicts the non-linearity between the pass-through of positive and negative shocks implied by the threshold value of promised utility: the difference between the average probability of leaving for shocks in the central bin and in the first bin is about 0.05, while the same difference with the last bin is half of the size. A similar argument holds for promised utility growth. Although, as described above, this non-linearity is crucial to generate the flatter pass-through of smaller positive shocks, it also implies that the pass-through of the large ones in the model is slightly smaller than the actual values.

below, also in non-recessions).

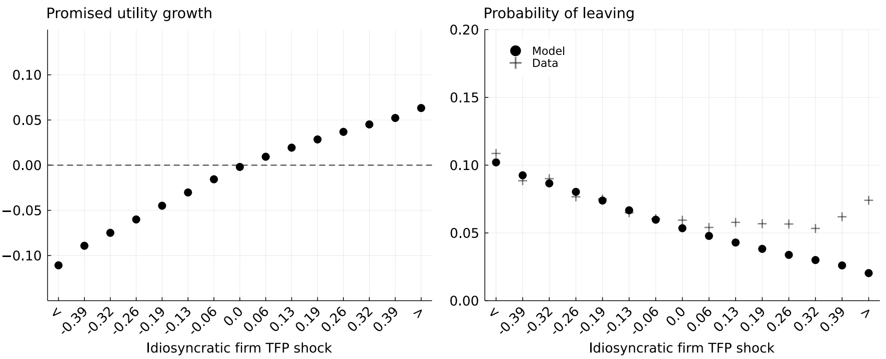


Figure 1.11: Promised utility growth and probability of leaving in steady state over the distribution of idiosyncratic firm productivity shocks.

By comparing the probability of leaving with the data, Figure 1.11 also provides supportive evidence in favor of the model's mechanism. Indeed, even though the empirical values for positive shocks are undershot, the probability of leaving is overall decreasing[39] across the support of firm's TFP shocks both in the model and the data. Remarkably, this outcome was obtained by targeting only the average job-to-job transition rate and not the full schedule.

The role of shocks' persistence. To further support the explanation provided above, Figure 1.12 compares the baseline pattern resulting from using the calibrated value of ρ against the ones obtained with different choices for the autocorrelation coefficient.[40] Indeed, the above discussion reveals a testable implication for the model mechanism: the lower the persistence of the shocks, the faster their reversion to the mean and, thus, to less extreme values of the match. In other words, the match's expected value after a large negative (positive) shock diminishes (increases) less for lower persistence values which should, in

[39]For large positive shocks it actually increases back, which suggests that other forces - not captured by the model - might be relevant in these cases.

[40]To ensure comparability across the different specifications, for each alternative value of ρ I change the normalization constant a in the production function to ensure that the average wage is the same as in the baseline economy.

turn, make the model mechanism less strong.

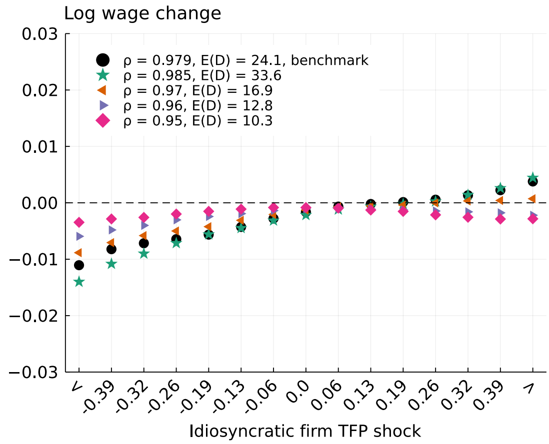


Figure 1.12: Average log wage change over the distribution of idiosyncratic firm productivity shocks for different values of the autocorrelation coefficient of firm productivity ρ. $\mathbb{E}(D)$ represents expected duration in quarters computed as $(1-\rho^2)^{-1}$.

Looking at the picture it is easy to see that this prediction is confirmed: the lower ρ is, the smaller the pass-through of shocks, especially if they are large, regardless of their sign. With lower persistence, large shocks have a smaller impact on the expected value of a match because they revert faster back to the mean. For negative shocks, this means that the firm is less willing to separate from the worker, which translates into smaller pass-through. On the other hand, for positive shocks, this means lower incentives to retain the worker, which also translates into smaller pass-through.

1.5.3 Business cycles

Building on the intuition behind the forces outlined in the previous section, I now use the full version of the model including the additional disaster state in recessions to describe the results for the pass-through of shocks over the business cycle.[41]

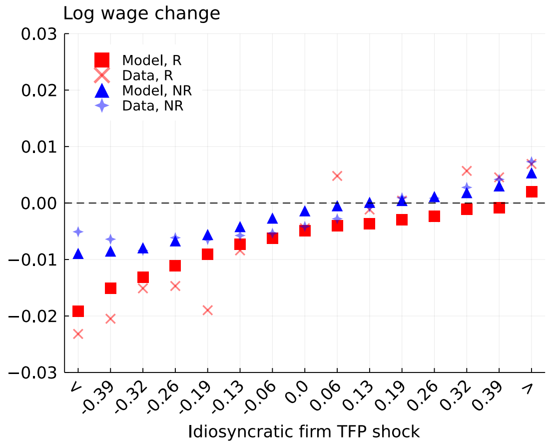


Figure 1.13: Average log wage change over the distribution of idiosyncratic firm productivity shocks, full model.

Figure 1.13 plots the average log wage change over the distribution of idiosyncratic firm shocks in recessions (squares) and non-recessions (triangles). Overall, the model is able to replicate the empirical patterns. The mechanisms behind the result are the same as the ones explained in the previous section. The upper bound of promised utility applies for positive shocks regardless of the business cycle, which explains the similar pattern for non-recessions and recessions. On the other hand, negative shocks are passed through much more in recessions because the possibility of disastrous events drastically reduces the expected value of the match which, in turn, boosts firms' incentives to dissolve the

[41] In this version of the model, b, a and γ are state-dependent. I rescale b by aggregate productivity in recessions and non-recessions and set a so that in the two aggregate states average idiosyncratic firm productivity is one. This ensures that unemployment benefits constitute the same share of average firm productivity. Finally, the values of γ are set so that the job-finding rate from unemployment in non-recessions is slightly above the value in the pooled case (since almost all years are defined as non-recessionary), 0.165, and the ratio between the model-implied job-finding rate from unemployment in recessions and non-recessions is about 0.88, which is in line with the corresponding figure in the Current Population Survey. This results in γ equal to 0.328 in R and 0.315 in NR.

match.

It is worth remarking that, while general-equilibrium effects would push the model mechanism in the direction of generating more pass-through of negative shocks in recessions without the need for the additional disaster state, they are quantitatively not sufficient to match the empirical trends. To see this, Figure 1.14 compares the results obtained in the baseline specification with those obtained from a counterfactual economy without the disaster state.[42] While a negative idiosyncratic shock of a given size - especially if large in absolute value - is associated with a larger wage decrease in downturns, because the calibrated length of recessions is relatively short, the value of the match is still too high for the firm to substantially increase its willingness to separate from the worker.

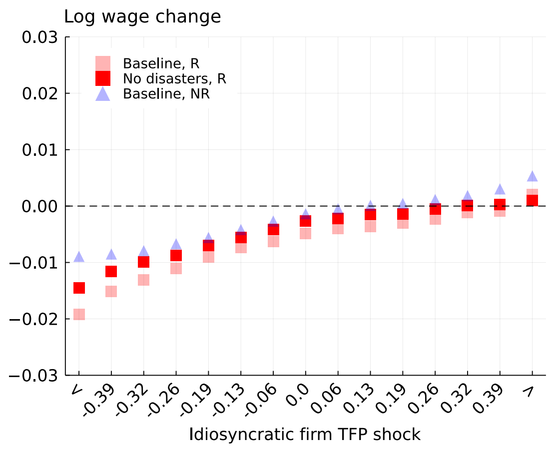


Figure 1.14: Average log wage change over the distribution of idiosyncratic firm productivity shocks, full model vs. model without firm disaster state.

[42] I also recalibrate the version of the model with business cycles but no disaster state following the same procedure described in footnote 41. In this case, γ is 0.302 in R and 0.309 in NR.

1.5.4 Implications for income risk over the business cycle

In a seminal paper, using administrative data from the US, Guvenen et al. (2014) show that the distribution of earnings changes features acyclical variance and procyclical skewness. These findings have later been extended to other countries with different institutional settings, namely Sweden, Germany, and France in a recent contribution by Busch et al. (2022). In Figure 1.5, I have shown that this also holds in my data when conditioning on my sample of stayers.

In principle, the mechanism outlined in the model can generate procyclical skewness. Indeed, if firms cut wages more in recessions upon receiving negative shocks and if the pass-through of positive shocks is much smaller and not state-dependent, this should generate more negative skewness when the economy is in a downturn.

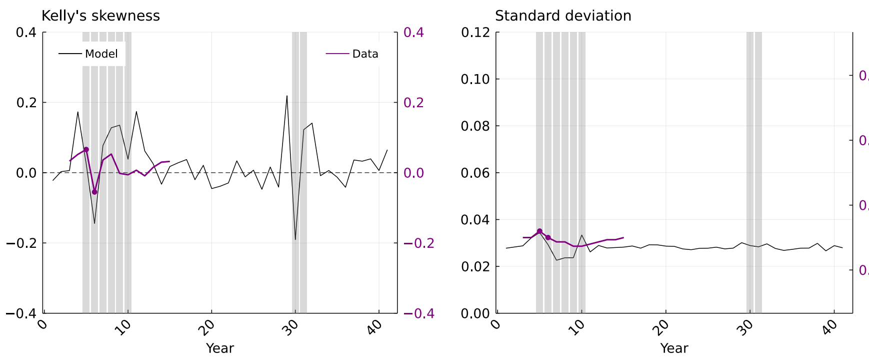


Figure 1.15: Kelly's skewness and standard deviation of the cross-sectional distribution of log wage changes over time computed from model simulations. Shaded areas correspond to recessions. The data series (right y-axis scale) is aligned so that the start of the recession in the data corresponds to the start of the recessionary period in the simulation. Recessionary periods in the data are indicated by dots.

To investigate the model's ability to generate the patterns in the data, Figure 1.15 plots Kelly's skewness and standard deviation of the cross-sectional distribution of log wage changes computed from model simulations. In line with the empirical findings, the distribution of log wage changes overall exhibits procyclical skewness and also acyclical standard deviation.[43]

This analysis supports the idea that it is rather the fact that more firms implement higher average wage cuts than the increased probability of extreme events that drives the behavior of the distribution in recessions. In addition, it also validates the model mechanism as an alternative explanation for business cycle trends in income risk and proves that wage pass-through is potentially one of their determinants.

1.5.5 Welfare cost of business cycles

Following the seminal approach in Lucas (1987), in this section I use the model to compute the welfare cost of business cycles. Let $\{c_{i,t}\}_{t=0}^{\infty}$ the consumption stream of individual i in the baseline economy without business cycles - with V its associated value function - and $\{\tilde{c}_{i,t}\}_{t=0}^{\infty}$ the stream in the alternative economy with business cycles - with $\tilde{V}$ its associated value function. The goal is to find how much consumption in the baseline economy the agent is willing to give up to avoid switching to the alternative economy. Mathematically, this means finding the value of λ that solves the following equation:

$$\mathbb{E}_0\left[\sum_{t=0}^{\infty}\beta^t u((1+\lambda)c_t)\right] = \mathbb{E}_0\left[\sum_{t=0}^{\infty}\beta^t u(\tilde{c}_t)\right] \tag{1.10}$$

Having the two utility values V and $\tilde{V}$ for each individual, one can solve for λ in (1.10) to obtain the consumption equivalent between the two series under scrutiny.[44] However, it is not straightforward in

[43]For readability, the figure reports only 40 years. Similar patterns hold over the whole simulation horizon.

[44]Given the utility function $u(x) = \frac{x^{1-\sigma}-1}{1-\sigma}$ and the definition of the value functions as discounted streams of future utility, from equation (1.10) one gets:

$$(1+\lambda)^{1-\sigma}\mathbb{E}_0\left[\sum_{t=0}^{\infty}\beta^t \frac{c_t^{1-\sigma}}{1-\sigma}\right] = \mathbb{E}_0\left[\sum_{t=0}^{\infty}\beta^t \frac{\tilde{c}_t^{1-\sigma}}{1-\sigma}\right] \iff \lambda = \left(\frac{\tilde{V}+\text{adj}}{V+\text{adj}}\right)^{\frac{1}{1-\sigma}} - 1$$

with $\text{adj} = [(1-\sigma)(1-\beta)]^{-1}$.

economies with high degrees of heterogeneity which utility values to choose (Krusell et al., 2010). This also applies to my model, where each individual has a specific value of promised utility evolving dynamically (i.e., V is a state) and it is not obvious which are the correct values to consider for the solution of (1.10).

I proceed as follows. For each value of firm productivity z, I get the distribution of promised utilities in the baseline and alternative specifications. Then, I compare the average promised utility in each quantile between these distributions and weigh the resulting λ by the share of agents in that position in the baseline model. For the baseline, I use the stationary distribution of promised utilities implied by the model without business cycles. For the alternative, I consider the distribution obtained in each period from simulating the model with aggregate uncertainty. The λs recovered through the procedure just described represent, therefore, the change in consumption between two individuals in the same relative position in the distributions of utilities in steady state and period t. In other words, how much consumption an individual in a given position in the steady state distribution of utilities is willing to give up in order not to find herself in the same relative position in the period t distribution. This approach, therefore, enables us to find a full cross-sectional distribution of λs in each period t, which can then be used for welfare comparisons. For instance, it is possible to compute the utilitarian consumption equivalent variation of the alternative economy by averaging the values of λ over all agents in the model in recessions, non-recessions, or all periods.

Table 1.5 reports the results. Overall, agents would be willing to give up 2.7 percent of their consumption to avoid aggregate uncertainty. While the numbers are relatively large compared to the literature, the lack of a saving technology and the presence of the disaster state make the effects of salary fluctuations particularly adverse in the model. Recessions are more costly than non-recessions, but the difference is not large as downturns are relatively short-lived in the used calibration.[45]

	All Periods	Recessions	Non-recessions
λ	-0.026	-0.027	-0.026

Table 1.5: Welfare costs of business cycles. Results refer to quarterly consumption.

1.6 Conclusion

Using Swedish administrative data covering the universe of firms, employment relationships, and workers, this paper has shown that earnings pass-through of idiosyncratic firm productivity shocks is asymmetric over the business cycle. Regardless of the state of the economy, a firm exposed to a positive shock passes it through mainly if it is sizable. On the other hand, firms are good insurance providers against negative shocks in normal times, but they pass them through much more in recessions. Specifically, compared to non-recessions, the earnings elasticity to negative shocks in downturns is more than three times higher. In monetary terms, a one standard deviation negative shock in recession implies - in the most conservative specification - a reduction in workers' annual salary by approximately 1,500 SEK (in 2020 terms). I have also documented that the share of firms experiencing mass layoffs upon receiving a negative idiosyncratic shock is about 7 percentage points higher in downturns. Combined with the fact that it takes around three years for them to return to positive employment growth, this shows that these shocks have larger and more persistent effects on the firm's profitability in recessions.

These empirical patterns have been rationalized in a directed search model of the labor market with on-the-job search, risk-averse

[45] As seen in section 1.5.2, the pass-through of idiosyncratic firm shocks is relatively limited in steady state. Thus, the estimated costs of business cycles are not far from the welfare costs of not providing full insurance to the workers.

workers, and firm commitment. The firm's trade-off between insuring the workers and making their on-the-job search choices aligned with the expected value of the match generates different degrees of pass-through across the distribution of idiosyncratic firm shocks. The non-state-dependent pass-through of positive shocks is delivered by the bound on the maximum utility that entering firms can offer to workers implied by the free-entry condition. Taking into account the more disastrous nature of negative shocks in recessions, instead, is crucial to generate their larger pass-through in downturns. As the model-generated wage growth distribution for stayers features procyclical skewness and acyclical variance, the theoretical framework also provides a new explanatory mechanism for recent empirical findings on business cycle trends in income risk. Welfare calculations reveal significant costs of business cycle fluctuations: to be compensated, workers would need to receive up to 2.7 percent of additional consumption.

The analysis presented in this work can be extended to several interesting avenues. First, while it is above the scope of this paper to provide a full analysis of the impact of the differences between my empirical strategy and the one adopted by Chan et al. (2020), my work shows that they seem relevant and deserve more investigation. In particular, it would be worth analyzing the effect of using earnings or wages as a measure of workers' salaries. Second, my analysis has considered the impact of pass-through on the average worker in a firm. It would be interesting to understand if the degree of insurance is heterogeneous also across different types of workers (e.g., white vs. blue collar). Third, this paper has focused on job market frictions to explain the empirical patterns found in the data. It would be relevant to understand if other mechanisms, such as firm labor market power (Chan et al., 2020), can also help rationalize them. Finally, the model provides a natural setting to study optimal public unemployment insurance over the business cycle in a context where firms already insure their workers. It is left to future research to investigate these important questions.

References

Arellano, M., Bonhomme, S., Vera, M. D., Hospido, L., and Wei, S. (2021). Income Risk Inequality: Evidence from Spanish Administrative Records. Working Papers 2136, Banco de España.

Azariadis, C. (1975). Implicit contracts and underemployment equilibria. *Journal of Political Economy*, 83(6):1183–1202.

Baily, M. N. (1974). Wages and Employment under Uncertain Demand. *Review of Economic Studies*, 41(1):37–50.

Balke, N. and Lamadon, T. (2022). Productivity shocks, long-term contracts, and earnings dynamics. *American Economic Review*, 112(7):2139–77.

Bloom, N., Floetotto, M., Jaimovich, N., Saporta-Eksten, I., and Terry, S. J. (2018). Really uncertain business cycles. *Econometrica*, 86(3):1031–1065.

Busch, C., Domeij, D., Guvenen, F., and Madera, R. (2022). Skewed idiosyncratic income risk over the business cycle: Sources and insurance. *American Economic Journal: Macroeconomics*, 14(2):207–42.

Carlsson, M., Clymo, A., and Joslin, K.-E. (2022). Dispersion over the business cycle: Passthrough, productivity, and demand. *Sveriges Riksbank Working Paper Series*, 414.

Carlsson, M., Messina, J., and Skans, O. N. (2016). Wage adjustment and productivity shocks. *The Economic Journal*, 126(595):1739–1773.

Chan, M., Salgado, S., and Xu, M. (2020). Heterogeneous passthrough from tfp to wages. *Working Paper*.

Friedrich, B., Laun, L., Meghir, C., and Pistaferri, L. (2019). Earnings dynamics and firm-level shocks. Working Paper 25786, National Bureau of Economic Research.

Guiso, L. and Pistaferri, L. (2020). The insurance role of the firm. *The Geneva Risk and Insurance Review*, 45(1):1–23.

Guiso, L., Pistaferri, L., and Schivardi, F. (2005). Insurance within the firm. *Journal of Political Economy*, 113(5):1054–1087.

Guvenen, F., Ozkan, S., and Song, J. (2014). The nature of countercyclical income risk. *Journal of Political Economy*, 122(3):621–660.

Hagedorn, M. and Manovskii, I. (2008). The cyclical behavior of equilibrium unemployment and vacancies revisited. *American Economic Review*, 98(4):1692–1706.

Hubmer, J. (2018). The job ladder and its implications for earnings risk. *Review of Economic Dynamics*, 29:172–194.

Juhn, C., McCue, K., Monti, H., and Pierce, B. (2018). Firm performance and the volatility of worker earnings. *Journal of Labor Economics*, 36(S1):S99–S131.

Knight, F. H. (1921). Risk, uncertainty and profit. *New York: Houghton Mifflin*.

Krueger, D., Mitman, K., and Perri, F. (2016). Macroeconomics and Household Heterogeneity. In Taylor, J. B. and Uhlig, H., editors, *Handbook of Macroeconomics*, volume 2 of *Handbook of Macroeconomics*, chapter 0, pages 843–921. Elsevier.

Krusell, P., Mukoyama, T., and Şahin, A. (2010). Labour-Market Matching with Precautionary Savings and Aggregate Fluctuations. *The Review of Economic Studies*, 77(4):1477–1507.

Krusell, P. and Smith, Jr., A. A. (1998). Income and wealth heterogeneity in the macroeconomy. *Journal of Political Economy*, 106(5):867–896.

Lucas, R. (1987). *Models of Business Cycles*. Oxford: Blackwell.

Menzio, G. and Shi, S. (2010). Block recursive equilibria for stochastic models of search on the job. *Journal of Economic Theory*, 145(4):1453 – 1494. Search Theory and Applications.

Menzio, G. and Shi, S. (2011). Efficient search on the job and the business cycle. *Journal of Political Economy*, 119(3):468–510.

National Mediation Office (2004). Avtalsrörelsen och lönebildningen 2004. medlingsinstitutets årsrapport. *Stockholm: Medlingsinstitutet*.

OECD (2021). *OECD Economic Surveys: Sweden 2021*.

Pagano, M. (2020). Risk sharing within the firm. *Working Paper*.

Rouwenhorst, K. G. (1995). Asset pricing implications of equilibrium business cycle models. In *Frontiers of Business Cycle Research*, pages 294–330. Princeton University Press.

Rudanko, L. (2009). Labor market dynamics under long-term wage contracting. *Journal of Monetary Economics*, 56(2):170–183.

Salgado, S., Guvenen, F., and Bloom, N. (2020). Skewed business cycles. *Working Paper*.

Schaal, E. (2017). Uncertainty and unemployment. *Econometrica*, 85(6):1675–1721.

Syverson, C. (2011). What determines productivity? *Journal of Economic Literature*, 49(2):326–65.

Topel, R. and Fredriksson, P. (2010). 3. wage determination and employment in sweden since the early 1990s: Wage formation in a new setting. In Freeman, R. B., Swedenborg, B., and Topel, R. H., editors, *Reforming the Welfare State*, pages 83–126. University of Chicago Press, Chicago.

von Wachter, T., Song, J., and Manchester, J. (2009). Long-term earnings losses due to mass layoffs during the 1982 recession: An analysis using u.s. administrative data from 1974 to 2004. *Working Paper*.

Appendices

1.A Additional figures

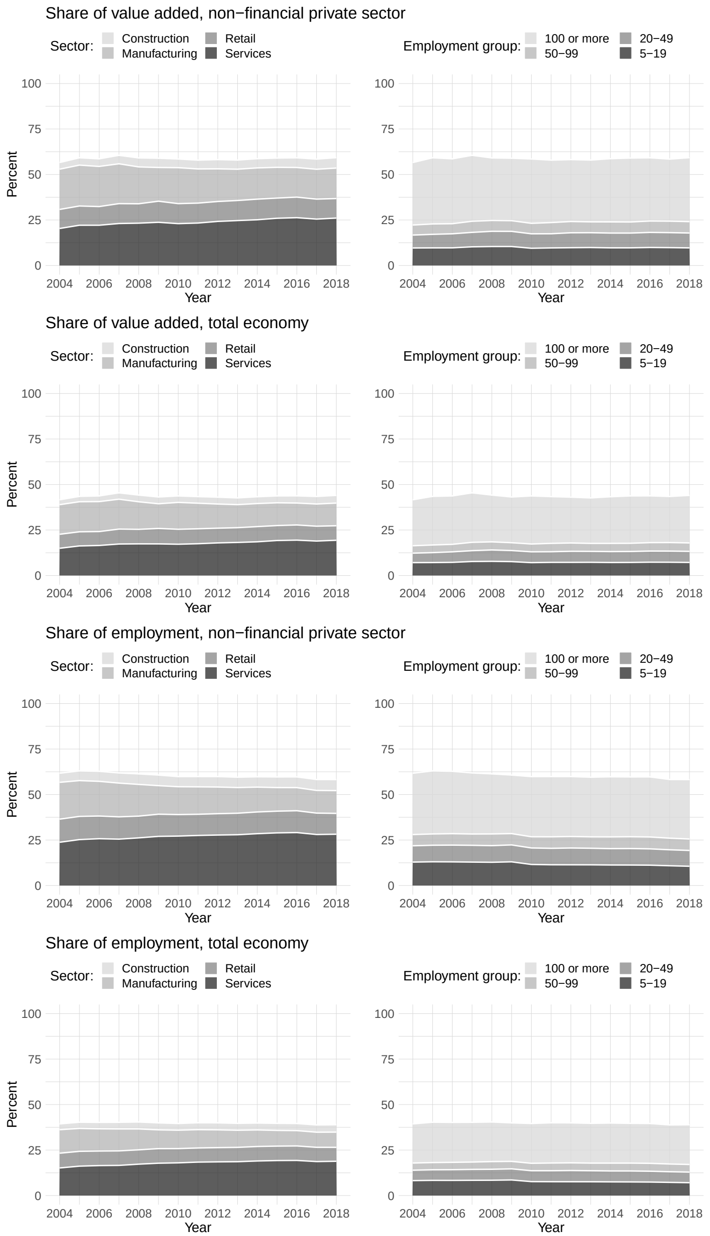


Figure 1.A.1: Share of aggregate value-added and employment in the non-financial private sector (first and second rows) and in the whole economy (third and fourth rows) by sector and employment group.

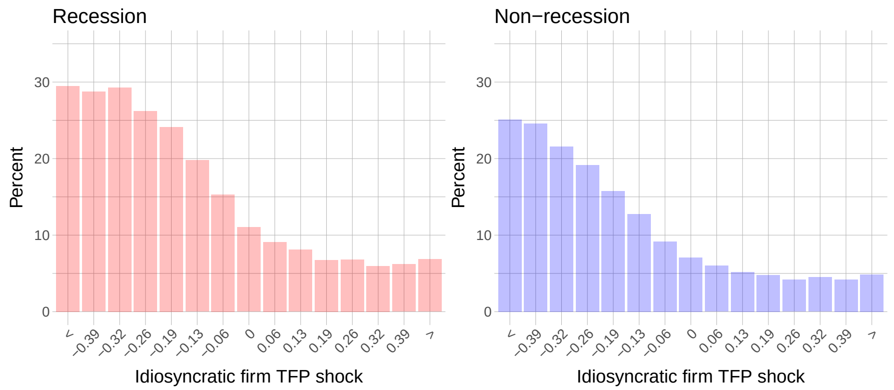


Figure 1.A.2: Share of firms experiencing mass layoffs over the distribution of idiosyncratic firm TFP shock, recessions and non-recessions.

1.B Data and institutional background

Structural Business Statistics (SBS). Because the unit of analysis considered in this paper is the firm[46], I use the dataset *SBS företag* which contains balance sheet, accounting, industry codes as defined by Statistics Sweden (SNI) and some other company information at the firm level. The time range of the sample I analyze is 1997-2018: before 1997 not all Swedish firms are included in the data and the last year available is 2018. I drop observations with missing firm identifier and firms having more than one entry per year. I also exclude observations without SNI code or for which I cannot convert the SNI code according to the 2007 definition.[47] Furthermore, I exclude firms that are classified as financial companies, firms with negative wage bill, firms with zero wage bill but at least one employee and firms with negative operating

[46] A firm in the dataset is a legal entity that can include one or more establishments. More firms can be owned by the same mother company.

[47] In the time range I consider SNI definitions changed twice. Specifically, the relevant SNI definition is SNI 1992 until 2002, SNI 2002 between 2003 and 2008 and SNI 2007 afterwards. In order to get a consistent definition of SNI codes across years, I extend the SNI 2007 definition back in time. To do this I use conversion files provided by Statistics Sweden and, when the conversion is not one-to-one, I use the most common transition in the data which I obtain by counting the number of transitions in the years around the switch in SNI definition.

expenses. Monetary variables are deflated using the CPI for Sweden.

Register-Based Labour Market Statistics (RAMS). On the firms' side, RAMS includes information on employment, juridic status, sector of activity and municipality from 1985 to 2019. The observation level is the establishment. I exclude observations with missing firm or establishment identifier and with missing information on the juridic form and municipality. Because municipality is observed at the establishment level and the unit relevant for my analysis is the firm, every time there is more than one municipality associated to a firm in a given year, I assign to the firm/year observation the municipality where most establishments are located and, if there is a tie, that of the establishment with higher number of employees. From the municipality codes I also recover the region in which the firm is based.

On the workers' side, RAMS includes, for every employed individual, the establishment and firm identifiers of all the workplaces she is employed at, the starting and ending month of each employment relationship, a measure of the salary received in each employment relationship and the occupational status of the worker (employee, self-employed, etc.) for the time period 1985-2019. I exclude observations with missing firm of worker identifier, those with missing information on the initial and final month of the employment relationship and I focus only on workers whose occupational status is of employee and who are employed at least six months in a year. Furthermore, if a worker has more than one employment relationship in a given year, for that year I classify that worker as employed at the firm where she receives the highest salary. Based on this, I construct a tenure indicator that keeps track of the number of years an employment relationship has been going on, a variable that counts the number of employees in each firm, and an annual measure of earnings by annualizing the salary received at the firm[48] where she has been classified to work at in that year as ex-

[48]This is done by dividing the total salary paid by the number of months in which the employment relationship has been active and multiplying the result by twelve.

plained before. Monetary variables are deflated using the CPI for Sweden.

Longitudinal Database on Education, Income and Employment (LOUISE). This dataset contains socio-economic and demographic information for each Swedish resident for the period 1985-2019. I set the lower bound of the time range to 1998 as before this year data on some civil statuses are not available. From this dataset I recover a household identifier, the civil status, gender, year of birth, number of children and education. I exclude observations with missing information on individual and household identifier, gender, year of birth, number of children, education or civil status and people with more than six children. I also merge education levels to get four levels (pre-secondary, secondary, post-secondary and post-graduate) and civil statuses to get four categories (singles, married, separated, survivors). Furthermore, I construct dummy variables capturing the period when a person becomes married, remarries or separates (by either divorcing or becoming a survivor). In addition, using the household identifier, I link married people to their partner and drop observations which are registered as married but for whom I did not get a match to find their partner.

Aggregate data. Nominal GDP, CPI (reference year 2020) and the unemployment rate are taken from the OECD Economic Outlook 110 (December 2021). Total and sectoral value-added and employment are taken from SCB's website: the data are quarterly so I aggregate them at the annual frequency by summing up the quarterly flow values in each year for value-added and and by averaging across the four quarters for employment. The OECD recession indicator ("OECD based Recession Indicators for Sweden from the Period following the Peak through the Trough, +1 or 0, Monthly, Not Seasonally Adjusted") was downloaded from FRED. Aggregation from monthly to annual is done as follows. First, I define a quarter as recessionary if the monthly indicator is one in all the three months of the quarter. Then, I define a year as recessionary if there are at least two consecutive recessionary quarters.

Institutional background. As pointed out by Guiso et al. (2005), in labor markets characterized by high levels of centralized wage bargaining there is much less leeway for firms to pass idiosyncratic shocks to workers' salaries. In this section, therefore, I briefly discuss some features of the Swedish institutional background in order to provide evidence that a substantial part of workers has at least part of their salaries determined at the firm level. The main source for this section is Topel and Fredriksson (2010).

While union membership rates in Sweden have been very high (around 80%) since the 70s, wage-setting institutions have changed quite considerably in the last fifty years. The highly centralized procedures in place during the 70s have indeed been gradually displaced in favor of decentralized agreements. Table 1.B.1 presents the wage agreement models present in the Swedish labor market in 2004 together with the percentage of employees in the private sector covered by each type of model.[49]

Model	Employees (%)
1. Local bargain without restrictions	7
2. Local bargain with a fallback	8
3. Local bargain with a fallback plus a guaranteed wage increase	16
4. Local wage frame without a guaranteed wage increase	12
5. Local wage frame with guarantee or a fallback regulating the guarantee	28
6. General pay increase plus local wage frame	18
7. General pay increase	11

Table 1.B.1: Percentage of employees in the private sector under different wage agreement models. *Source:* Topel and Fredriksson (2010) based on National Mediation Office (2004).

[49] A fallback means that the central agreement specifies a general wage increase that comes into operation should the local parties not agree. A guaranteed wage increase means that each individual is guaranteed a wage increase of a certain amount of Swedish krona (SEK). A local wage frame means that the local parties are given a total wage increase but can decide on the distribution of that increase over individuals.

The table reveals that, while there is substantial heterogeneity in the distribution of wage agreement models - with at the exterma 11% of workers subject only to the central agreement and 7% to local bargain without restrictions - 89% of workers are into agreements in which potentially at least part of the salary is determined at the local level and 19% in local agreements in which a wage increase is not guaranteed. Therefore, there is room for firm performance to be reflected in workers' wages.

1.C Robustness

Alternative definitions of business cycles. In order to check that the main empirical finding is not determined by the business cycle episodes classification, I have also experimented with two alternative definitions: (i) defining recession 2008-2009 and non-recession 2014-2015 (ii) defining recessions using the OECD recession indicator for Sweden.[50] Figure 1.C.3 plots GDP growth and the unemployment rate when using these alternative definitions.

[50]See Appendix 1.B for more details on the OECD indicator.

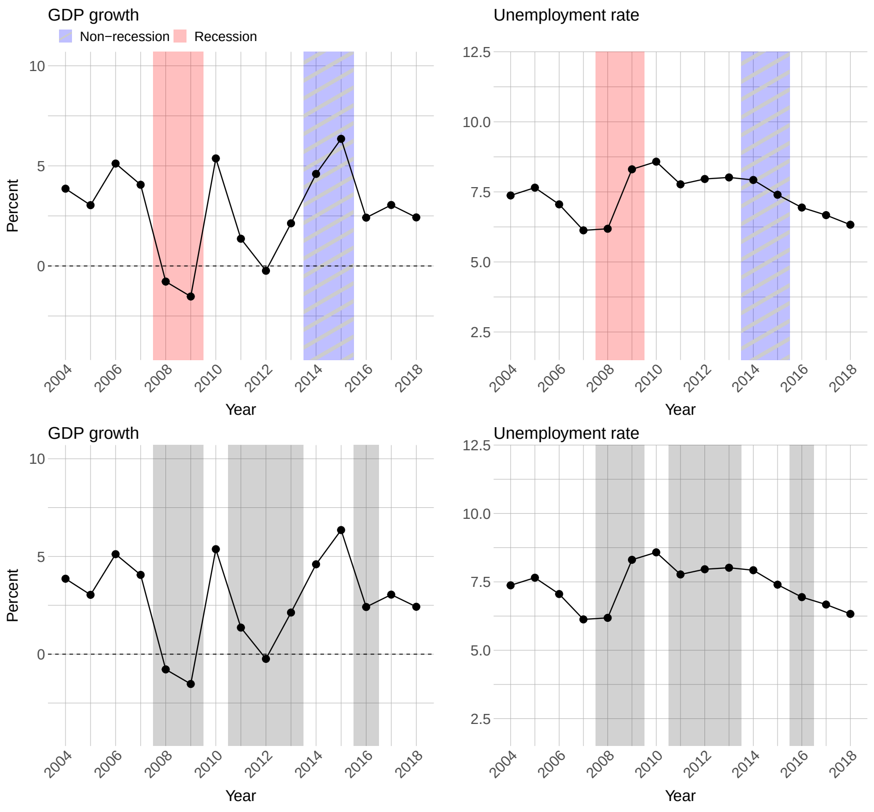


Figure 1.C.3: GDP growth and unemployment rate in Sweden and alternative business cycle definitions. *Top:* recession 2008-2009, non-recession 2014-2015. *Bottom:* recessions and non-recessions defined according to OECD recession indicator.

Figure 1.C.4 reports the earnings pass-through graphs for these alternative definitions. Even though when using the OECD indicator the distinction for negative shocks between recessions and non-recessions becomes slightly less stark, overall the two pictures clearly indicate robustness of the main empirical result.

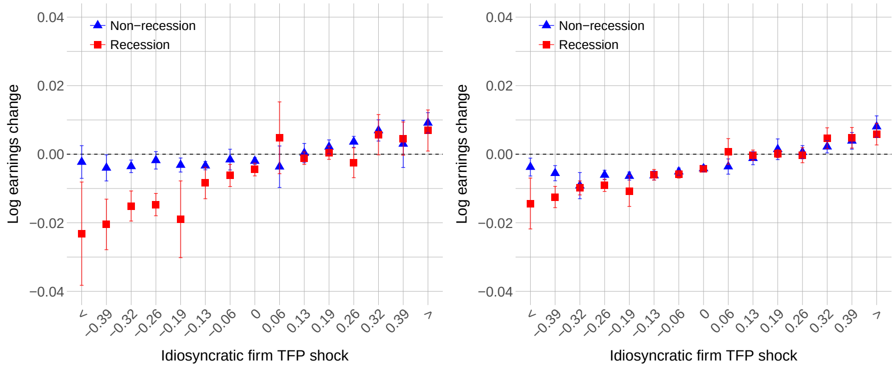


Figure 1.C.4: Average residual log earnings change over bins of idiosyncratic firm shocks for alternative business cycle definitions, recessions (squares) and non-recessions (triangles). *Left*: recession 2008-2009, non-recession 2014-2015. *Right*: recessions and non-recessions defined according to OECD recession indicator. Confidence intervals are at the 90 percent level and standard errors are clustered at the firm level. The first bin contains all firms below the -0.39 bin and the last all firms above the 0.39 bin.

Alternative measure of firm productivity. In order to check the robustness of the results to the measure of firm productivity used in the baseline analysis - log value-added per worker - in this section I replicate the main empirical finding using an alternative definition. Following - among others (see Syverson, 2011, for a review) - Bloom et al. (2018) and Salgado et al. (2020), under the assumption that the firm production function is Cobb-Douglas, it is possible to recover a measure of firm productivity as the residual of the following equation:

$$\log Y_{jst} = a_{st}^N \log N_{jst} + a_{st}^K \log K_{jst} + z_{jst}$$

where Y_{jst} is value-added of firm j operating in sector s at time t, N_{jst} and K_{jst} are, respectively, measures of the labor and capital inputs used in the production process, a_{st}^N and a_{st}^K the factor shares and the residual, z_{jst}, is firm productivity.

Practically, to recover z with this approach, I proceed as follows. First, I compute the labor share a_{st}^N as the ratio between the total wage bill and total value-added in each two-digit sector s and each year t.[51][52] Using the identity $a^K_{st} = 1 - a^N_{st}$ I obtain the capital share. Then, using the wage bill as measure of N, total fixed assets as measure of K and the measure of value-added already available in the data as Y, I recover z_{jst} from the above formula. Finally, I run regression (1.1) using this new measure of productivity and recover the new shocks ν_{jt}. Figure 1.C.5 presents the results. Overall, even though the magnitude of wage changes is smaller and positive large shocks do not seem to be passed through much, the trends reported in the baseline are robust.

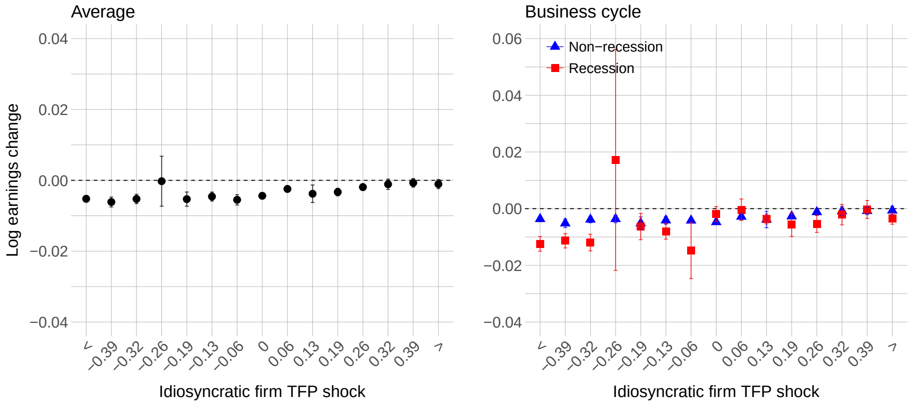


Figure 1.C.5: Average residual log earnings change over bins of idiosyncratic firm shocks for alternative firm productivity definition. *Left*: all years. *Right*: recessions (squares) and non-recessions (triangles). Confidence intervals are at the 90 percent level and standard errors are clustered at the firm level. The first bin contains all firms below the -0.39 bin and the last all firms above the 0.39 bin.

Heterogeneity. In this section, I investigate possible heterogeneities in the results documented in the empirical section of the paper.

Figure 1.C.6 plots the residual log earnings changes for firms of different sizes (number of employees between 5-19, 20-49, 50-99 and 100+). Overall, the patterns are still there for all the four groups, even though

[51] All the monetary variables used are in 2020 SEK.
[52] I exclude the sector-year cells with less than 100 firms.

the business cycle asymmetry for negative shocks is slightly larger for larger firms.

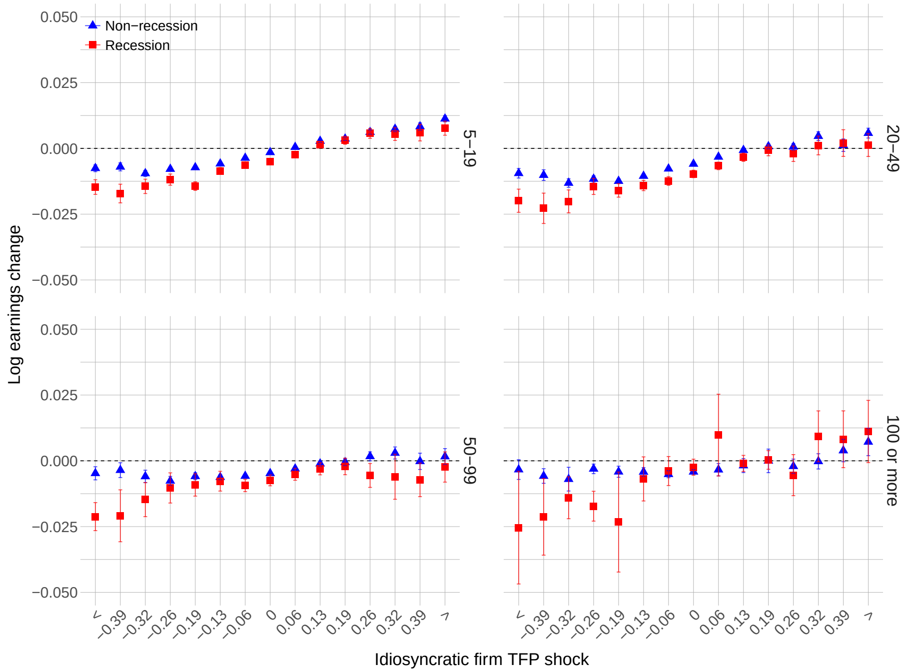


Figure 1.C.6: Average residual log earnings change over bins of idiosyncratic firm shocks, firms' size heterogeneity. Confidence intervals are at the 90 percent level and standard errors are clustered at the firm level. The first bin contains all firms below the -0.39 bin and the last all firms above the 0.39 bin.

Figure 1.C.7 plots the same thing for workers in different age bands (20-33, 34-47, and 46-60). In general, the empirical trends still hold across all groups.

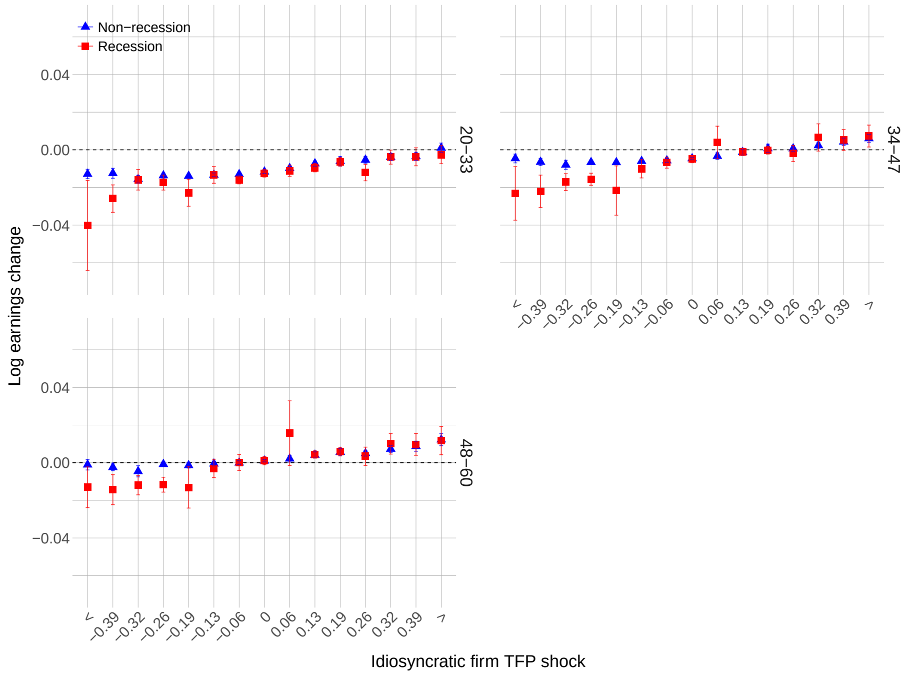


Figure 1.C.7: Average residual log earnings change over bins of idiosyncratic firm shocks, workers' age heterogeneity. Confidence intervals are at the 90 percent level and standard errors are clustered at the firm level. The first bin contains all firms below the -0.39 bin and the last all firms above the 0.39 bin.

Figure 1.C.8 investigates heterogeneities in workers' tenure (I divide workers in three groups, with *low* indicating shorter tenure and *high* longer tenure). Again, the empirical patterns are overall still there, even though the negative shocks' asymmetry is stronger for more tenured workers.

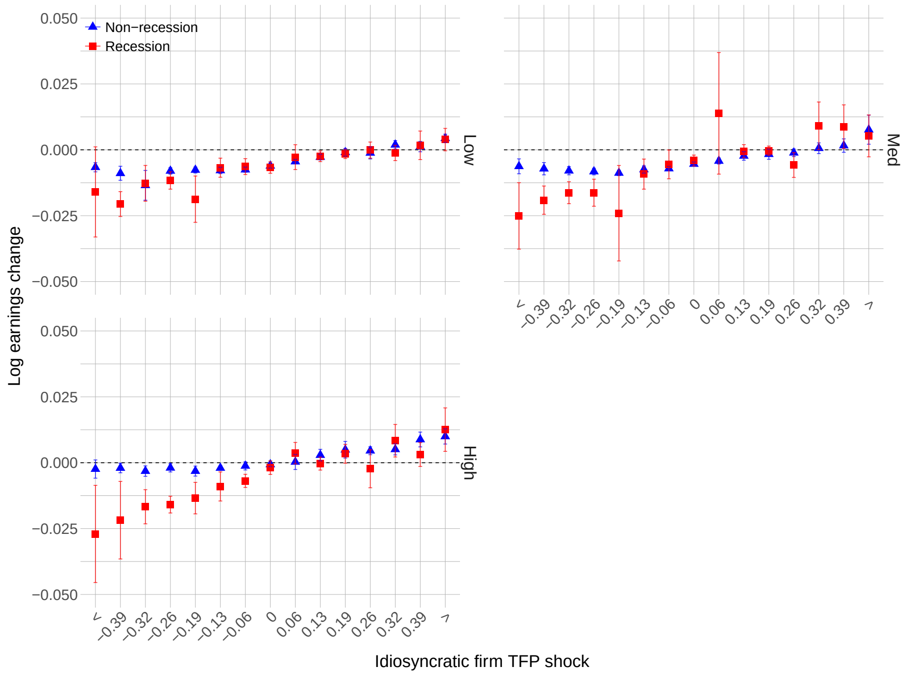


Figure 1.C.8: Average residual log earnings change over bins of idiosyncratic firm shocks, workers' tenure heterogeneity. Confidence intervals are at the 90 percent level and standard errors are clustered at the firm level. The first bin contains all firms below the -0.39 bin and the last all firms above the 0.39 bin.

1.D Block recursivity of the equilibrium

As per definition 1.3.2, proving that the equilibrium of the model economy is block recursive requires to show that the equilibrium objects $\{\theta, R, x^*, U, J, w, \hat{V}\}$ depend on the aggregate state ψ just through the aggregate state of nature s and not through the distribution of workers g_u, g_e. Following closely Menzio and Shi (2010), the proof will consist in showing that each of the equilibrium objects and the operator T for updating the firm value satisfy this property.

Let $\mathcal{J}(Y \times Z \times X)$ be the set of firm value functions $J : Y \times Z \times X \to \mathbb{R}$ satisfying properties (J1)-(J3) described in Menzio and Shi (2010). Take an arbitrary function $J \in \mathcal{J}$. For all $(\psi, x) \in \Psi \times X$ satisfying $J(y, z_0, x) \geq k$ the market tightness implied by the equilibrium condition (1.7) is $q^{-1}\left(\frac{k}{J(y,z_0,x)}\right)$ and for all $(\psi, x) \in \Psi \times X$ such that $J(y, z_0, x) < k$ tightness is zero. Given the properties of J, the former condition is satisfied if and only if $x \leq \tilde{x}(y)$, where $\tilde{x}(y)$ solves $J(y, z_0, x) = k$. Therefore, market tightness can be summarized as follows:

$$\theta(y, x) = \begin{cases} q^{-1}\left(\frac{k}{J(y,z_0,x)}\right) & \text{if} \quad x \leq \tilde{x}(y) \\ 0 & \text{otherwise} \end{cases}$$

As it is clear from the above expression, the market tightness function θ depends on the aggregate state ψ only through aggregate productivity y. Intuitively, since the firm value does not depend on the distributions g_u, g_e and the cost of creating a vacancy is constant, the probability of filling a vacancy - and thus tightness - is independent of the distribution of workers. In addition, the properties of J and q imply that tightness is decreasing in x. Intuitively, it is easier to attract a worker - and thus to fill a vacancy - by promising her higher utility.

Turning to the return to search function R, note that, given θ, for all $(\psi, x) \in \Psi \times X$ the objective function $\max_{x \in X} p\left(\theta(y, x)\right)(x - V)$ in (1.3) depends on the aggregate state just through aggregate productivity. In addition, the choice set X does not depend on the aggregate state of the economy. Thus, the optimal search decision m and the return to search functions R depend on the aggregate state just through y and not through g_u, g_e. Intuitively, this follows from the fact that both the two terms in the objective function - the job finding probability and the return - are independent of the employment status of other workers.

In turn, letting $\mathcal{U}(Y)$ be the set of unemployment values $U : Y \to \mathbb{R}$ starting from any initial guess $U \in \mathcal{U}$ the operator in equation (1.4) is a contraction mapping in which the updated value of unemployment in each iteration depends on the aggregate state just through y and not on the distribution of workers because both the unemployment benefit and the return to search are independent of the distribution of workers.

With all the above it is now possible to construct the updated value of the firm $\tilde{J}$ by plugging the elements just described in the optimization problem of the firm:

$$\tilde{J}(y,z,V) = \max_{w,\{\hat{V}(\hat{y},\hat{z})\}_{\forall(\hat{y},\hat{z})}} f(y,z) - w \\ + \beta(1-\delta)\mathbb{E}\left[\left(1-\lambda_e\tilde{p}\left(\hat{y},\hat{V}(\hat{y},\hat{z})\right)\right) J(\hat{y},\hat{z},\hat{V}(\hat{y},\hat{z}))\right]$$

subject to:

$$V = u(w) + \beta\mathbb{E}\left[\delta U(\hat{y}) + (1-\delta)\left(\hat{V}(\hat{y},\hat{z}) + \lambda_e R(\hat{y},\hat{V}(\hat{y},\hat{z}))\right)\right]$$

Because the objective function above and the choice sets depend on aggregate productivity y but not on the distribution of workers, the same holds true for the updated value of the firm. Intuitively, this is a consequence of the fact that the match output, the match survival probability and the continuation value are depend on y but not on g_u, g_e. Repeating the same argument by using as new firm value $\tilde{J}$ it is possible to conclude that the operator defining the equilibrium value of the firm T also depends on y but not on g_u, g_e, which completes the proof.

1.E Numerical solution

Discretization and grids construction. I have already specified in Section 1.4 of the main text the processes for aggregate productivity y and for firm TFP z, so in this section I will describe just the construction of the grid of promised utilities and the discretization of the AR(1) process governing the evolution of z.

Grid for promised utility. Following Menzio and Shi (2010), the lowest and highest values for the grid of promised utilities are set as follows:

$$\underline{x} = \frac{u(b)}{1-\beta} - \varepsilon_x, \quad \overline{x} = \frac{u(f(\overline{y},\overline{z}))}{1-\beta} + \varepsilon_x$$

The grid of promised utilities is then a grid of N_x points between these two values with spacing parameter spacing_x.

Discretization idiosyncratic firm productivity $\tilde{z}$. This is simply discretized with the method proposed by Rouwenhorst (1995) with N_z points.

Table 1.E.2 summarizes the choices for the numerical parameters.

Parameter	Value	Description
Panel A: numerical parameters for model solution		
N_x	401	Number of points promised utility grid
N_z	29	Number of points idiosyncratic firm TFP
ε_x	1	Value for promised utility grid extrema
spacing_x	1.5	Spacing parameter promised utility grid
Panel B: numerical parameters for model simulation		
T_{sim}	10000	Number of simulated periods
N_{sim}	2000	Number of simulated agents
$T_{\text{dis,sim}}$	1000	Number of periods to discard for moments computation
Panel C: numerical parameters for SMM estimation		
N_{glo}	100	Number of points to evaluate in global stage, estimation $\tilde{z}$
	5000	Number of points to evaluate in global stage, estimation ζ
N_{loc}	10	Number of points to evaluate in local stage, estimation $\tilde{z}$
	20	Number of points to evaluate in local stage, estimation ζ
N_{eco}	5	Number of economies to simulate
T_{cal}	5000	Number of simulated time periods
T_{dis}	500	Number of periods to discard for moments computation
N_{cal}	1500	Number of simulated firms

Table 1.E.2: Numerical parameters.

Solution procedure. Solving the full model is a very challenging exercise. In each iteration towards the equilibrium - given the current guess - there are four steps to perform: (1) solve for market tightness (2) solve workers' search problem (3) solve for the problem of the unemployed workers (4) solve the problem of the firm and update. I will outline the solution of each subproblem and then lay out the full solution algorithm. Note that, in order to have a good starting guess when solving the full model, I first solve a version of the model without on the job search. The four steps are present in each of the two cases, but I will highlight differences between the solution of the full model and the simpler model in each stage when required.[53]

Step 1: market tightness. Let i indicate the current iteration number and $J^{(\mathrm{i}-1)}$ the current initial guess for the value of the firm. From the free-entry condition (1.7) we know that in all visited markets the relation must hold with equality, so that, rearranging we can find tightness as follows:

$$\theta^{(\mathrm{i})}(s,x) = q^{-1}\left(\frac{k}{J^{(\mathrm{i}-1)}(s,z_0,x)}\right) \tag{1.E.1}$$

Thus, for all $s \in S$ and for all $x \in X$:

- compute the ratio $\frac{k}{J^{(\mathrm{i}-1)}(s,z_0,x)}$;
- because $q \in [0,1]$, if the ratio is not in this interval then use complementary slackness and set $\theta^{(\mathrm{i})}(s,x) = 0$;
- otherwise, compute tightness using equation (1.E.1).

Furthermore, by inserting the obtained values into the function for the job finding probability, I obtain $p^{(\mathrm{i})}(\theta)$. Because it will be needed afterwards, at this step I also compute for all $s \in S$ the exact last market $x^{\text{switch}}(s)$ where it is profitable for firms to post vacancies. This market solves:

$$J(s, z_0, x^{\text{switch}}(s)) = k$$

which I solve with a numerical solver.

Step 2: workers' search problem. Workers' search problem is summarized by the return to search function (1.3) and its corresponding first order condition (1.8). There are three relevant intervals for all $s \in S$: (i) for all $x \in [\underline{x}, V]$ the return to search is lower or equal than

[53] I will assume that block recursivity holds, so I will replace the aggregate state ψ with s in the description below to ease up the explanation.

zero (ii) for all $x \in (V, x^{switch}(s))$ the return to search is strictly positive and strictly concave in x (iii) $x \in [x^{switch}(s), \overline{x}]$ the return to search is zero. Therefore, the solution must be in the interval in (ii) and because of strict concavity the first order condition is necessary and sufficient to characterize the maximum. Thus, for all $s \in S$ and for all $V \in X$:

- if $V \geq x^{\text{switch}}(s)$ then set $x^{*(\text{i})}(s, V) = V$;
- else, solve with a numerical solver the first order condition (1.8) with respect to $x^{*(\text{i})}(s, V)$ in the interval $x \in (V, x^{\text{switch}}(s))$;
- insert $x^{*(\text{i})}(s, V)$ into the job finding probability previously computed and into the expression for the return of search to get $\tilde{p}^{(\text{i})}(s, V)$ and $R^{(\text{i})}(s, V)$.[54]

The policy for x^* has to be computed very precisely. However, because the derivative of the job finding probability tends to minus infinity for values of x approaching x^{switch}, and because I do not have a closed form expression for it but rather have to compute it numerically, solving for the FOC close to this point is potentially problematic. To overcome this issue, I approximate[55] the job finding probability p - which is a function of the promised utility values - as follows:

$$a_0 + a_1(V - \underline{x})^2 \tag{1.E.2}$$

which allows me to compute the derivative in closed form.

Step 3: unemployed workers' problem. Solving the problem of the unemployed amounts to solve the fixed point problem defined by (1.4). Specifically, using as starting guess for the unemployment value $\underline{x}$, I proceed as follows:

- for all $s \in S$, compute the new guess for the value of unemployment using equation (1.4);

[54]Computing $\tilde{p}$ is not strictly necessary to solve the simpler problem without on the job search.

[55]A similar approximation procedure is used by Balke and Lamadon (2022).

- if the maximum of the absolute value of difference between the old and new guess is lower than the convergence threshold then define the last obtained value as the solution for the unemployment value $U^{(\mathrm{i})}$;

- otherwise, keep iterating until convergence.

Step 4: firms' problem and updating. The firms' problem can be divided into three substages: (i) finding the optimal promised utilities in each future state of the world (ii) finding the optimal wages (iii) updating the firm value. In the simpler model without on the job search, note that the first order condition with respect to the future promised utility (1.9) becomes very simple: the right hand side becomes zero due to the lack of search frictions and the optimal policy prescribes constant wages. Therefore $W(\hat{\xi}) = V$ for all V, ξ and, substituting this into the promise keeping constraint I get $V = u(w) + \beta\delta\mathbb{E}\left[U(\hat{\psi})\right] + \beta(1-\delta)V$ which can easily be solved for w for all V, ξ. One can then immediately substitute these optimal policies to get the updated firm value.

The firms' problem with on the job search is instead solved as follows. Let $w^{(\mathrm{i}-1)}$ be the current guess for the wage policy (as starting guess I use the wage policy obtained from the simpler problem without on the job search). Then:

- for all $\xi \in S \times Z$ and all $V \in X$, using the current guess for the wage policy, for all $\hat{\xi} \in S \times Z$ use equation (1.9) to solve for $W^{(\mathrm{i})}(\hat{\xi})$;

- replace the $W(\hat{\xi})$ obtained into the promise keeping constraint (1.6), which then can be solved for the new guess for $w^{(\mathrm{i})}$ for all $\xi \in S \times Z$ and all $V \in X$;

- for all $\xi \in S \times Z$ and all $V \in X$ get the new guess of the firm value $J^{(\mathrm{i})}$ with equation (1.5) using the policies $w^{(\mathrm{i})}$ and $W^{(\mathrm{i})}$ just computed.

The policy for W has to be computed very precisely. However, this is very hard for the same reasons described above for the optimal search choice. Therefore, I follow the same approach as before and approximate $\tilde{p}$ with the functional form 1.E.2, which makes things easier as now I have a closed form expression for the derivative. Finally, because also the new guess for the firm value has to be computed very precisely, I approximate that too with the same functional form 1.E.2, which also allows me to get a smooth version of the updated wage policy $w^{(\mathrm{i})}$ through the envelope condition $-\frac{\partial J}{\partial \hat{V}} = \frac{1}{u'(w)}$, which I use a starting guess for the next iteration.

Solution algorithm.

I. Set initial guesses:

- for the simple model without on the job search, set the firm value as a strictly concave function of the grid for promised utilities;
- for the full model set the initial firm value and the initial guess for the wage policy as those obtained from solving the simple model without on the job search;

II. Solve for market tightness as described in Step 1;

III. Solve the workers' search problem as described in Step 2;

IV. Solve the unemployed workers' problem as described in Step 3;

V. Solve the firms' problem and update as described in Step 4;

VI. Compute the difference between the value of the firm from the original guess and the update at all grid points below the last open market. If the maximum - in absolute value - of such difference is lower than the tolerance threshold then exit the loop, otherwise go back to point II with the new updated guesses.

1.F Estimation

This section describes how I estimate the parameters governing the exogenous stochastic processes not already obtained from closed form expressions described in the calibration section. The numerical parameters chosen for the estimation procedure are reported in Table 1.E.2.

Let θ be the vector of parameters that has to be estimated. θ is chosen to minimize the following objective function:

$$\min_{\theta} \hat{m}(\theta)' W \hat{m}(\theta) \tag{1.F.3}$$

where $\hat{m}(\theta)$ is a vector of moments that depends on the parameters to be estimated and W is a weighting matrix. The procedure involves a global and a local stage. In the global stage I compute the value of the objective function for N_{glo} combination of points for the elements of the vector θ. The combinations correspond to the first N_{glo} of a Sobol sequence. At the end of the global stage, the best - in the sense of providing the lowest values of the objective function - N_{loc} points pass to the local stage. In the local stage, for each of the N_{loc} points, equation (1.F.3) is solved for the minimum using the Nelder-Mead algorithm with starting guess each of such points. The minimum is then the vector of parameters among the N_{loc} local points that returns the lowest value of the objective function.

To estimate the parameters governing idiosyncratic firm productivity, I simulate its exogenous process for T_{cal} periods and for N_{cal} firms. I then aggregate the model-generated data at the annual frequency (I take as annual measure the values in the last quarter and I define a year as recessionary if there are at least two consecutive quarters in which the economy is in downturn) and run regression (1.1) on the simulated data to recover the model-generated shocks ν. Finally, I match the model generated moments and the moments in the data. I discard the first T_{dis} points from moments computation and, in order to smooth the surface of the objective function, I simulate the process for

N_{eco} economies and average moments across them. Letting m indicate a generic moment, $\hat{m}(\theta)$ is defined as follows:

$$\hat{m}(\theta) = \frac{m_{\text{data}} - m_{\text{simulated}}(\theta)}{m_{\text{data}}} \tag{1.F.4}$$

The weighting matrix W is a diagonal unitary matrix when ζ is inactive and otherwise its values related to the share of firms experiencing mass layoffs take the value $0.15/2$, those related to the duration of the disaster state and the share R/NR of firms exiting the disaster state take the value $0.7/2$ and Kelley's skewness takes the value 0.15. The targeted moments are described in detail in Section 1.4 of the main text.

Chapter 2

Inferring income properties from portfolio choices

I am grateful to Per Krusell, Paolo Sodini, Kurt Mitman and Kieran Larkin for advice, discussions and support. I also wish to thank Timo Boppart, Tobias Broer, Richard Foltyn, José-Elías Gallegos, John Hassler, Stefan Hinkelmann, Philipp Hochmuth, Markus Kondziella, John Kramer, Claire Thurwächter, Zoltán Rácz, Roine Vestman, Joshua Weiss and participants at the IIES Macro Group for helpful feedback and comments.

2.1 Introduction

Income risk is a crucial determinant of agents' choices and it is therefore of paramount importance to understand its nature in order to correctly evaluate the effect of economic policies.

While the literature on this topic is vast, two main hypotheses have emerged: according to one, income shocks are very persistent and agents face similar life-cycle profiles - *Restricted Income Profiles* (RIP); according to the other, income shocks are not very persistent and life-cycle profiles are individual-specific - *Heterogeneous Income Profiles* (HIP). Income data alone do not easily allow us to empirically discern which of the two is right. As shown by Guvenen (2009), identification is achieved by looking at income covariances at far apart ages in the life cycle. These covariances are however hard to compute due to natural attrition. A further problem is the known issue that the researcher's information set is a subset of the agent's. Several papers have, therefore, taken another approach in order to tackle this question: because diverse types of income risk imply different economic choices, by looking at the latter the researcher can infer properties of the income process. For instance, Blundell et al. (2008), Guvenen (2007) and Guvenen and Smith (2014) consider consumption/saving choices.

This paper contributes to this literature by looking at another decision of the agent that is determined by income risk, namely portfolio choice. While the canonical portfolio choice model does not, in general, admit a closed form solution for the share of wealth detained in risky assets when labor income is not deterministic, several papers (e.g., Cocco et al., 2005; Chang et al., 2018b; Catherine, 2021) have analyzed the portfolio choice implications of income risk with numerically solved models. The main finding is that the properties of the income process - and especially higher order moments - are relevant for the portfolio allocation. Starting from this result, this paper investigates to what extent portfolio choice models can be used to understand the nature of the income process. Specifically, i) I estimate HIP and RIP versions of a rich stochastic process including cyclical skewness of income shocks, ii) I use them as inputs to solve a state of the art portfolio choice model, and iii) I find identifying restrictions from differences in model outcomes.

My findings are the following. First, cyclical skewness needs to be included in the income process in order to correctly estimate the amount of risk deriving from the persistence of the shocks. Indeed, in a model without cyclical skewness, HIP overestimates the share of risk attributed to heterogeneity in life-cycle profiles and underestimates the share deriving from persistence.

Second, because including cyclical skewness results in similar estimates for the persistence of the shocks for both HIP and RIP, I find that the profiles of the mean and variance of consumption over the life cycle are very much alike. In light of this, these two schedules, which have previously been considered by the literature (e.g., Guvenen, 2007), do not have a strong identification power to infer income properties.

Third, I find that HIP and RIP imply different average life-cycle profiles for participation and conditional risky share. This is due to the fact that agents' choices for these two quantities are different in the two models depending on their expected income growth rate during working life. Intuitively, the latter is more similar across agents in RIP and more heterogeneous in HIP. As a consequence, agents in HIP expecting higher lifetime income choose a lower risky share if human capital is risky, especially when they do not have enough wealth to self-insure. Specifically, due to the effect of cyclical skewness on the riskiness of human capital, the HIP process implies much less heterogeneity in participation rates across people with different income growth rates, and a "butterfly pattern" for the conditional risky share. The latter means that people with high income growth rates choose lower conditional risky shares in young ages compared to individuals with low growth rates, and they catch up at around 40 years old when the order of this pattern is reversed.

Comparing the model-generated profiles and their empirical equivalents using Swedish administrative data, I find that the latter provides slightly stronger support for the RIP than the HIP hypothesis, especially because no evidence of a "butterfly pattern" was found.

Related literature. First, this paper is related to the area of research that studies the nature and properties of the income process. Given the central role of the latter in economics and finance, this literature is unsurprisingly vast (e.g., Lillard and Willis, 1978; MaCurdy, 1982; Meghir and Pistaferri, 2004; Storesletten et al., 2004; Guvenen, 2009; Guvenen et al., 2014, 2021; De Nardi et al., 2019; Busch et al., 2022; Arellano et al., 2022). My contribution here is showing that once cyclical skewness of labor income shocks is taken into account in the estimation of the income process, the estimated persistence coefficient of the shocks is not much different between HIP and RIP.

Within this literature, a series of papers (e.g., Hall and Mishkin, 1982; Deaton and Paxson, 1994; Blundell and Preston, 1998; Blundell et al., 2008; Kaplan and Violante, 2010; Arellano et al., 2017) have used revealed choices to make inference on the properties of the income process. The main idea is that, because the latter has a relevant impact on agents' decisions, it should be possible to infer something about the income risk individuals face by looking at their choices. As mentioned, this approach also helps solve the well-known issue that the econometrician's information set is a subset of the agent's. Nevertheless, these papers have mainly looked at consumption choices. My contribution is, thus, to use this method on another decision that is influenced by income risk, namely portfolio choice. In using a structural model to make inference, my paper is most closely related to the approach followed by Guvenen (2007) and Guvenen and Smith (2014).[1]

Furthermore, this work relates to the literature on portfolio choice models and income risk. That income risk has an impact on portfolio choices has been already investigated empitically and theoretically

[1]Among these two papers, it is linked especially to the former as my objective is also to disentangle HIP and RIP.

(e.g., Cocco et al., 2005; Gomes and Michaelides, 2005; Benzoni et al., 2007; Storesletten et al., 2007; Gomes and Michaelides, 2008; Huggett and Kaplan, 2016; Chang et al., 2018b; Catherine, 2021; Merton, 1969; Fagereng et al., 2017; Chang et al., 2018a). A recent contribution by Catherine (2021) in this area has shown that modelling the procyclicality of the skewness of labor income shocks improves the fit with portfolio choices in the data. While keeping this channel in my model, this paper is the first to study the effects of cyclical skewness of labor income shocks and of different assumptions on the heterogeneity of life-cycle income profiles jointly.

Finally, as the framework in this paper is used to make inference on the properties of the income process, this paper is also connected to the emerging household finance literature in (Calvet et al., 2021) using portfolio choice models for structural estimation.

Structure of the paper. The paper is structured as follows. Section 2.2 presents the model, Section 2.3 deals with model estimation and calibration, Section 2.4 presents the results and Section 2.5 concludes.

2.2 Model

The model is a refined version of the canonical portfolio choice model by Cocco et al. (2005). Specifically, I build on recent work from Catherine (2021), who shows that including cyclical skewness of persistent income shocks is crucial to match well portfolio choices in the data. Compared to Catherine (2021), my framework additionally includes heterogeneity in life-cycle income profiles: in my model the constant and the coefficient of the age trend in the individual-specific part of income are state variables.[2]

[2]Catherine (2021) estimates the parameters governing the income process assuming that individuals have different life-cycle income profiles but does not include this heterogeneity when solving the model.

Preferences. Let t denote age. Each agent enters into the model at age T_{start}, lives at maximum until age T and works until retirement age K. Each agent has Epstein-Zin preferences, similar to the forms specified in Inkmann et al. (2010) and Gomes and Michaelides (2005):

$$U_{i,t} = \left\{ (1-\delta) c_{i,t}^{1-\psi} + \delta \mathbb{E}_t \left[p_t U_{i,t+1}^{1-\gamma} + b(1-p_t) a_{i,t+1}^{1-\gamma} \right]^{\frac{1-\psi}{1-\gamma}} \right\}^{\frac{1}{1-\psi}} \tag{2.1}$$

where $c_{i,t}$ is individual i's consumption at t, $a_{i,t}$ is individual i's wealth at t, δ is the discount factor, p_t is the probability of being alive at $t+1$ conditional on being alive at t, γ is the coefficient of RRA, $1/\psi$ the EIS and b determines the strength of the bequest motive. Because there is a bequest motive, the terminal condition for the recursion is:

$$U_{i,T+1} = b a_{i,T+1}^{1-\gamma} \tag{2.2}$$

Financial assets. Agents can invest in two financial assets, one risky with time-varying gross return R_{t+1} and one safe with constant gross return R_f. Letting small letters indicate log returns, r_{t+1} is given by the following expression:

$$r_{t+1} = r_{1,t+1} + r_{2,t+1} - \kappa_m \tag{2.3}$$

The effective return an individual gets by investing in the risky asset is the sum of two systematic components, one co-varying with labor market conditions (r_1) and the other that does not (r_2), and is net of a management cost κ_m, that is thus paid conditional on holding the risky asset. The systematic components are modelled as in Catherine (2021). Specifically, to take into account stock market crashes, r_1 is modelled as a mixture of Normals:

$$r_{1,t+1} = \begin{cases} \underline{r}_{1,t+1} \overset{i.i.d.}{\sim} \mathcal{N}\left(\underline{\mu}_r, \sigma_{r_1}^2\right) & \text{w.p.} \quad p_r \\ \overline{r}_{1,t+1} \overset{i.i.d.}{\sim} \mathcal{N}\left(\overline{\mu}_r, \sigma_{r_1}^2\right) & \text{w.p.} \quad 1-p_r \end{cases} \tag{2.4}$$

Without loss of generality, it is possible to interpret p_r as the probability of stock market crashes and $\underline{\mu}_r$ the expected log return during crashing periods. Similarly, $1 - p_r$ is the probability of normal periods and $\overline{\mu}_r$ the average log return during normal periods. r_2 is, instead, a simple Normal shock:

$$r_{2,t+1} \overset{i.i.d.}{\sim} \mathcal{N}\left(0, \sigma_{r_2}^2\right)$$

Finally, investing in the risky asset is subject to a fixed participation cost κ_f that is paid every time the agent chooses to hold the risky asset. In addition, agents are allowed to borrow up to a borrowing limit on their total savings proportional to the exogenously set parameter $\overline{s}$. The repayment rate per unit of borrowing is set to be equal to the risk free rate.

Income process. I closely follow Catherine (2021) for specifying the labor income process. Let $L_{i,t}$ be individual i's real income. The logarithm of $L_{i,t}$ is the sum of an aggregate income component w_t and of an idiosyncratic component $y_{i,t}$:

$$\log\left(L_{i,t}\right) = w_t + y_{i,t} \tag{2.5}$$

The aggregate component follows a random walk with drift, driven by shocks to the market return through a parameter λ_{rw}:

$$w_t = g + w_{t-1} + \lambda_{rw} r_{1,t} + \phi_t \tag{2.6}$$

where $\phi_t \overset{i.i.d.}{\sim} \mathcal{N}\left(0, \sigma_\phi^2\right)$.

The idiosyncratic component is the sum of a deterministic life-cycle component $\bar{f}(t)$, of a persistent component $z_{i,t}$ and of a transitory component $\nu_{i,t}$:

$$y_{i,t} = f_{i,t} + z_{i,t} + \nu_{i,t} \tag{2.7}$$

The persistent component is an AR(1) process:

$$z_{i,t} = \rho z_{i,t-1} + \varepsilon_{i,t} \tag{2.8}$$

with innovations drawn from a mixture of Normals:

$$\varepsilon_{i,t} = \begin{cases} \underline{\varepsilon}_{i,t} \overset{i.i.d.}{\sim} \mathcal{N}\left(\underline{\mu}_{\varepsilon,t}, \underline{\sigma}_{\varepsilon}^2\right) & \text{w.p.} \quad p_\varepsilon \\ \overline{\varepsilon}_{i,t} \overset{i.i.d.}{\sim} \mathcal{N}\left(\overline{\mu}_{\varepsilon,t}, \overline{\sigma}_{\varepsilon}^2\right) & \text{w.p.} \quad 1 - p_\varepsilon \end{cases} \tag{2.9}$$

Without loss of generality, it is possible to interpret p_ε as the probability of tail events and $\underline{\mu}_{\varepsilon,t}$, $\underline{\sigma}_{\varepsilon,t}$ the expected value and standard deviation of persistent income shocks during tail events, respectively. A similar interpretation holds for the parameters governing the distribution of normal events. To match the cyclicality of skewness, $\underline{\mu}_{\varepsilon,t}$ is defined as:

$$\underline{\mu}_{\varepsilon,t} = \mu_\varepsilon + \lambda_{\varepsilon w}(w_t - w_{t-1}) \tag{2.10}$$

Thus, tail events imply on average higher persistent shocks during expansions and vice versa during recessions. In addition, because persistent idiosyncratic shocks have zero mean, it must hold:

$$p_\varepsilon \underline{\mu}_{\varepsilon,t} + (1 - p_\varepsilon)\overline{\mu}_{\varepsilon,t} = 0 \tag{2.11}$$

The transitory shock is a pure innovation following a Normal distribution, whose variance depends on whether the persistent shock was drawn from the tail distribution or not:

$$\nu_{i,t} = \begin{cases} \underline{\nu}_{i,t} \overset{i.i.d.}{\sim} \mathcal{N}\left(0, \underline{\sigma}_{\nu}^2\right) & \text{if} \quad \varepsilon_{i,t} = \underline{\varepsilon}_{i,t} \\ \overline{\nu}_{i,t} \overset{i.i.d.}{\sim} \mathcal{N}\left(0, \overline{\sigma}_{\nu}^2\right) & \text{if} \quad \varepsilon_{i,t} = \overline{\varepsilon}_{i,t} \end{cases} \tag{2.12}$$

Finally, the deterministic component is specified differently for the HIP vs. RIP processes. Mathematically, the two specifications are summarized as follows:

$$f_{i,t}^{\text{model}} = \begin{cases} \bar{f}(t) + \alpha_i + \beta_i t & \text{if model} = \text{HIP} \quad (2.13) \\ \bar{f}(t) & \text{if model} = \text{RIP} \quad (2.14) \end{cases}$$

where $\bar{f}$ is a function of experience that is common to all individuals and to both models. In the HIP case, α_i and β_i are drawn from an i.i.d.

bivariate Normal distribution with zero mean, variances σ_α^2, σ_β^2 and covariance $\sigma_{\alpha\beta}$. The RIP model, instead, only includes the function of experience and has no individual-specific components. Therefore, the RIP process is a restricted versions of the HIP specification with $\beta_i = \alpha_i = 0$ for all i.

Payroll taxes. As in Catherine (2021), working agents are subject to a 12.4% payroll tax on income up to a maximum taxable amount set by the Social Security Administration (SSA), which is roughly 2.5 times the average wage index.[3] Specifically, the total tax paid is:

$$T_{i,t} = .124 \cdot \min\{L_{i,t}, 2.5 \cdot e^{w_t}\} \tag{2.15}$$

Retirement income. It is common practice in standard portfolio choice models (e.g., Cocco et al., 2005) to assume that retirement income is a constant fraction of the income received in the last working period. However, in this model, large shocks in the last working period would have enormous effects on total retirement income, which is counterfactual. To solve this issue, I follow Catherine (2021) and assume that retirement income depends on the whole income history of an agent.[4] Let $\bar{L}_{i,t}$ keep track of average income as follows:

$$\bar{L}_{i,t} = \frac{1}{t - T_{\text{start}} + 1} \sum_{j=T_{\text{start}}}^{t} \min\{e^{y_{i,j}}, 2.5\} \tag{2.16}$$

Then, during retirement ($t > K$) agents have the following income stream:

$$\log(L_{i,t}) = \log(\zeta_{i,K}) + w_K \tag{2.17}$$

that is, retired agents enjoy a fraction ζ of the average wage index at the time of retirement. Following Catherine (2021), $\zeta_{i,K}$ depends on aver-

[3] In 2010 the average wage index was 41673.83 USD and the maximum taxable amount 106800 USD.

[4] The downside is that this implies taking care of another state variable in the maximization problem.

age income at retirement $\bar{L}_{i,K}$ as follows:

$$\zeta_{i,K} = \begin{cases} 0.9 \cdot \bar{L}_{i,K} & \text{if} \quad \bar{L}_{i,K} < 0.2 \\ 0.116 + 0.32 \cdot \bar{L}_{i,K} & \text{if} \quad 0.2 \leq \bar{L}_{i,K} < 1 \\ 0.286 + 0.15 \cdot \bar{L}_{i,K} & \text{if} \quad \bar{L}_{i,K} \geq 1 \end{cases} \tag{2.18}$$

Safety net. In order not to overstate the real level of income risk that agents face, it is important to model some traits of the welfare system. Therefore, following Catherine (2021), I model the Supplemental Nutrition Assistance Program and, for retired individuals, the Supplemental Security Income Program.

Only individuals with low wealth can participate in these programs. Specifically, only agents with less than about 2000 USD, which is roughly 5% of the average wage index. After retirement, eligible individuals receive supplemental income such that their total income reaches at least 20% of the average wage index. Before retirement, eligible individuals with earnings below 20% of the wage index receive benefits equal to 6% of the wage index minus 30% of their earnings. Mathematically:

$$N_{i,t} = \begin{cases} \max\{.06 \cdot e^{w_t} - .3 \cdot L_{i,t}, 0\} & \text{if} \quad a_{i,t} < .05 \cdot e^{w_t}, \\ & \qquad L_{i,t} < 0.2 \cdot e^{w_t}, t \leq K \\ \max\{.2 \cdot e^{w_t} - L_{i,t}, 0\} & \text{if} \quad a_{i,t} < .05 \cdot e^{w_t}, t > K \end{cases} \tag{2.19}$$

where $a_{i,t}$ is wealth as defined below.

The optimization problem. Let $\Xi_{i,t} = (\alpha_i, \beta_i, a_{i,t}, z_{i,t}, w_t, \bar{L}_{i,t})$ denote the state variable during working life and $\Xi^{\mathrm{R}}_{i,t} = (a_{i,t}, \bar{L}_{i,K}, w_K)$ the state variable during retirement.Also, let $R^e_{t+1} := \exp(r_{t+1}) - R_f$ denote the excess return and $L^d_{i,t} = L_{i,t} - T_{i,t} + N_{i,t}$ denote disposable income. Agent i chooses consumption $c_{i,t}$, savings $s_{i,t}$, a dummy variable $F_{i,t}$ equal to one in case she decides to hold risky assets and, conditional on participation, the share of savings invested in risky assets $\xi_{i,t}$ to maximize:

$$V_{i,t}(\Xi_{i,t}) = \max_{\xi_{i,t}, c_{i,t}, s_{i,t}, F_{i,t}} \Big\{ (1-\delta) c_{i,t}^{1-\psi} + \delta \mathbb{E}_t \left[p_t V_{i,t+1}^{1-\gamma}(\Xi_{i,t+1}) + b(1-p_t) a_{i,t+1}^{1-\gamma} \right]^{\frac{1-\psi}{1-\gamma}} \Big\}^{\frac{1}{1-\psi}} \tag{2.20}$$

subject to:

$$a_{i,t} = c_{i,t} + s_{i,t} + F_{i,t}\kappa_f \cdot \exp\{w_t\} \tag{2.21}$$

$$a_{i,t+1} = \left[R_f + \xi_{i,t} R^e_{t+1} \right] s_{i,t} + L^d_{i,t+1} \tag{2.22}$$

$$s_{i,t} \geq \bar{s} \cdot \exp\{w_t + \bar{f}(t)\} \tag{2.23}$$

and to the equations governing the exogenous processes outlined before for the different specifications governing the deterministic part of idiosyncratic income. During retirement, the agent's problem becomes:

$$V_{i,t}(\Xi^{\mathrm{R}}_{i,t}) = \max_{\xi_{i,t}, c_{i,t}, s_{i,t}, F_{i,t}} \Big\{ (1-\delta) c_{i,t}^{1-\psi} + \delta \mathbb{E}_t \left[p_t V_{i,t+1}^{1-\gamma}(\Xi^{\mathrm{R}}_{i,t+1}) + b(1-p_t) a_{i,t+1}^{1-\gamma} \right]^{\frac{1-\psi}{1-\gamma}} \Big\}^{\frac{1}{1-\psi}} \tag{2.24}$$

subject to the constraints (2.21), (2.22), (2.23), to the terminal condition (2.2) and to the equations governing the exogenous processes during retirement.[5] Appendix 2.C describes in detail how the model is solved numerically.

[5]In the last period of working life, that is when $t = K$, the problem is the same as under retirement, except that $V_{i,K}$ is defined on the state for the last period of working life $\Xi_{i,K}$.

2.3 Estimation and calibration

In this section, I describe the estimation of the parameters governing the exogenous stochastic processes for aggregate shocks, for individual income in the HIP and RIP cases and how I calibrate the remaining parameters in the model. The estimation procedure for aggregate shocks and individual income processes is described in Appendix 2.D. More details on the data used can be found in Appendix 2.B.

Aggregate processes. Estimation of the exogenous processes governing the market return and the aggregate part of income requires to estimate eight parameters: $\underline{\mu}_r, \overline{\mu}_r, \sigma_{r_1}, \sigma_{r_2}, p_r, \sigma_\phi, g, \lambda_{rw}$. I target mean, standard deviation, third and fourth standardized moments (skewness and kurtosis) of log yearly SP500 returns and aggregate wage log growth and the correlation between these two series.[6] The time sample for the returns is 1900-2019 and the one for aggregate wage growth is 1979-2011.[7] Table 2.1 reports the parameter estimates and the moments in the data and in the model. Overall the match is quite satisfactory.

Panel A: estimated parameters

Stock market returns					Aggregate income		
$\underline{\mu}_r$	$\overline{\mu}_r$	σ_{r_1}	σ_{r_2}	p_r	g	λ_{rw}	σ_ϕ
-0.242	0.114	0.074	0.114	0.138	0.008	0.170	0.016

Panel B: moments

	Log returns				Aggregate income shocks				
	Mean	SD	Skew	Kurt	Mean	SD	Skew	Kurt	Corr
Data	0.064	0.183	-0.635	3.352	0.019	0.029	-0.767	3.773	0.658
Model	0.065	0.183	-0.636	3.485	0.019	0.029	-0.771	3.631	0.654

Table 2.1: Estimated parameters for the stochastic processes governing macroeconomic aggregates and moments in the data vs. model.

[6] Specifically, the correlation is between the log return in year $t-1$ and aggregate wage log growth in year t.

[7] Data for returns are obtained from Robert Shiller's website and for aggregate wage growth from Guvenen et al. (2014). See Appendix 2.B for more details.

Individual income process. Estimation of the stochastic process governing individual income requires finding the values of eight parameters common to both specifications, namely $p_\varepsilon, \mu_\varepsilon, \lambda_{\varepsilon w}, \underline{\sigma}_\varepsilon, \overline{\sigma}_\varepsilon, \underline{\sigma}_\nu, \overline{\sigma}_\nu, \rho$, and two additional parameters for the HIP process, namely $\sigma_\beta, \sigma_{\alpha\beta}$. I target the time series between 1978 and 2010 of the standard deviation of log earnings growth at the one and five year horizons, Kelly's skewness of log earnings growth at the one, three and five year horizons and the within-cohort variance of log earnings for ages between 25 and 60.[8] The first set of moments is used to estimate the parameters of the persistent and transitory shocks and the second for the income profiles. In total I have 155 time-series moments and 36 within-cohort variances for a total of 191 moments. I perform SMM estimation, assuming that the economy is hit by the same aggregate wage shocks found in the data between 1944 and 2011.[9][10] with diagonal weighting matrix with all entries equal to one.

Table 2.2 reports the parameter estimates and Figure 2.1 plots the targeted moments in the different model specifications and in the data. Overall the match is satisfactory.

[8] I take the values for standard deviation and for Kelly's skewness from Guvenen et al. (2014) and the within-cohort variances from Guvenen et al. (2021). See Appendix 2.B for more details.

[9] I use data from Emmanuel Saez before 1979 and from Guvenen et al. (2014) afterwards. See Appendix 2.B for more details.

[10] I simulate the income histories of 68 cohorts, the first starting in 1944 and the last in 2011 assuming that the persistent component is zero at the beginning. Note that this implies having a constant age structure between 1979 and 2011. The same procedure is used in Guvenen et al. (2014) and Catherine (2021).

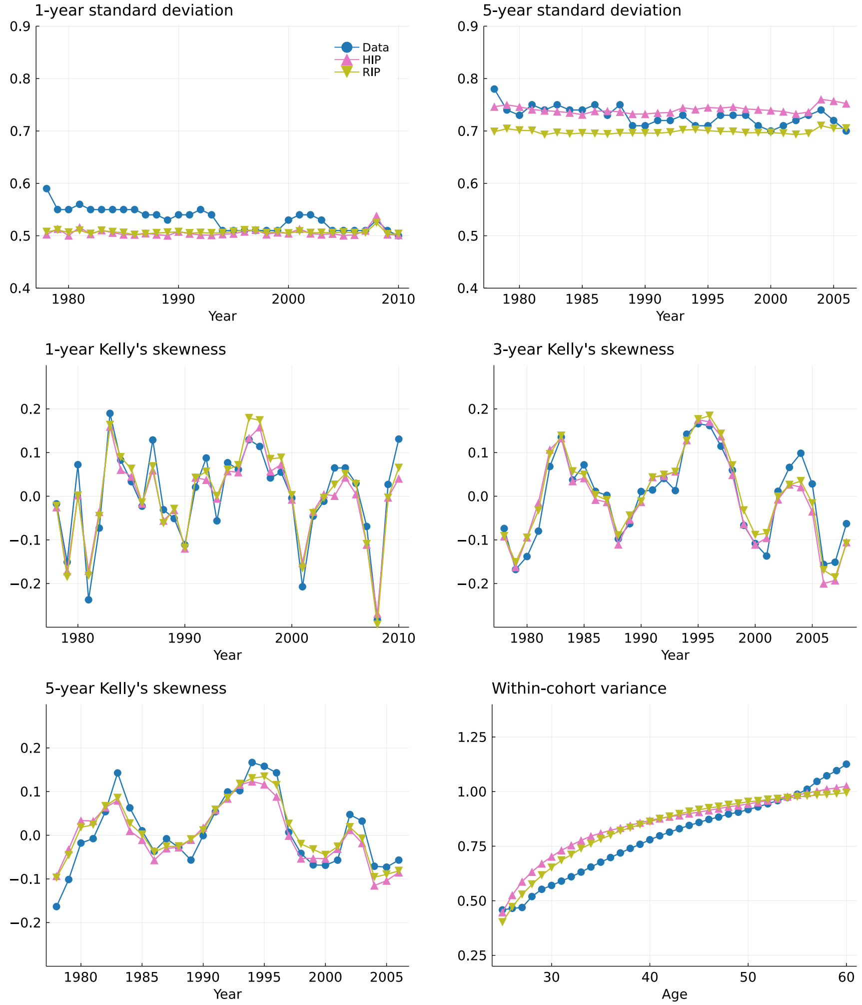


Figure 2.1: Individual income moments, comparison between different model specifications and data.

The table also reports estimated parameters when cyclical skewness is turned off.[11] In this case, I match only the standard deviation of

[11]In the model without cyclical skewness the shocks ε and ν are i.i.d. Normals with zero mean and standard deviations σ_ε and σ_ν, respectively. For readability these two values are reported in the same column of the tail-event standard deviation in the full specification.

log earnings growth at the one and five year horizons and the within-cohort variance of log earnings for ages between 25 and 60. Figure 2.A.1 compares the moments generated by the estimated specifications without cyclical skewness against the data.

	Persistent						Transitory		Income profiles		
	μ_ε	$\lambda_{\varepsilon w}$	$\underline{\sigma}_\varepsilon$	$\overline{\sigma}_\varepsilon$	ρ	p_ε	$\underline{\sigma}_\nu$	$\overline{\sigma}_\nu$	σ_α	σ_β	$\sigma_{\alpha\beta}$
HIP	-0.095	4.486	0.657	0.046	0.929	0.192	0.603	0.124	0.393	0.015	-0.005
RIP	-0.068	3.500	0.611	0.048	0.967	0.163	0.744	0.085			
Without cyclical skewness											
HIP			0.223		0.830		0.357		0.665	0.022	-0.012
RIP			0.216		0.977		0.352				

Table 2.2: Estimated parameters for the stochastic processes governing individual income.

In addition to reporting the parameter estimates, Table 2.2 shows an important result: in the model with cyclical skewness, the autocorrelation coefficient ρ is quite similar between HIP and RIP, while it is very different between them when cyclical skewness is turned off. The explanation is the following. Without cyclical skewness, the model attributes the total income risk faced by the agents on the persistence coefficient (ρ), on the variances of the transitory and persistent shocks ($\sigma_\nu, \sigma_\varepsilon$) and, in the HIP case, to the heterogeneity of the parameters governing the life-cycle income profiles ($\sigma_\alpha, \sigma_\beta, \sigma_{\alpha\beta}$). As it is evident by looking at the estimates, in the RIP case the risk due to heterogeneity in life-cycle profiles is almost entirely loaded on the autocorrelation coefficient, since the estimates for the variances of the shocks are very similar. However, by comparing these results with the higher part of the table, it is possible to see that, when taking into account skewness, the HIP model attributes a lower part of risk to the life-cycle profile component and a higher part to persistence. The estimate for ρ in the RIP case, instead, is lower but does not change much. The conclusion is, therefore, that including skewness is crucial not to overestimate the role played by the heterogeneity in life-cycle profiles in explaining income risk against the role played by persistence of the shocks.

Average life-cycle income profile. I model the average life-cycle income profile $\bar{f}(t)$ as a 3rd degree polynomial of age. I fit the polynomial on average log income by age reported by Guvenen et al. (2021) for the age range 25-60 from which, following Catherine (2021), I subtract a 15% average income tax.[12][13] Table 2.3 reports the estimated coefficients while Figure 2.A.2 plots the estimated profile against the data.

Other parameters. The remaining parameters are chosen as reported in Table 2.3. Agents' starting age is set to 23, they retire at 64 and they die with certainty at age 100.[14] Survival probabilities are taken from US life tables provided by SSA. I set the discount factor δ to 0.927 and the elasticity of intertemporal substitution $1/\psi$ to 0.336 as in Catherine (2021). I choose a standard value in the literature, namely 5 for risk aversion γ. The parameter governing the strenght of the bequest motive b is set to 2.5 as in Gomes and Michaelides (2005). As in Catherine (2021) the risk-free rate is set to 2% and the management fee to 1%. I set the fixed participation cost to 1.5% of the average wage. Finally, the borrowing limit is set to zero.

[12] I do not normalize age, that is, the coefficients are estimated using the age range 25-60.

[13] The values for ages outside this range are then extrapolated from the estimated model.

[14] The reference for these parameters is Catherine (2021), but these values are standard in the literature.

Parameter	**Value**	**Description**	**Source/Target**
Life-cycle			
T_{start}	23	Initial age	Catherine (2021)
K	64	Retirement age	Catherine (2021)
T	100	Maximum life span	Catherine (2021)
p_t	US life tables	Survival probabilities	SSA
Constant	-5.087	Average life-cycle income polynomial coefficients	Average log income by age from Guvenen et al. (2021)
t	0.249		
$t^2/10$	-0.042		
$t^3/100$	0.002		
Preferences			
δ	0.927	Discount factor	Catherine (2021)
γ	5	Risk aversion	Preset
ψ	$1/0.336$	Inverse EIS	Catherine (2021)
b	2.5	Bequest motive	Gomes and Michaelides (2005)
Financial markets and borrowing limit			
r_f	0.02	Risk-free rate	Catherine (2021)
κ_m	0.01	Management fee	Catherine (2021)
κ_f	0.015	Fixed participation cost	Preset
$\bar{s}$	0	Borrowing limit	Preset

Table 2.3: Other parameters.

2.4 Results

2.4.1 Life-cycle profiles

Figure 2.2 plots the average simulated life-cycle profiles for participation, conditional and unconditional risky share and wealth during working life for three-year age groups.[15][16] I compare the model-generated patterns against data from the eleven waves of the Survey of Consumer Finances (1989-2019). More details on the data and the computation of the empirical profiles can be found in Appendix 2.B.

[15] I use savings $s_{i,t}$ as measure of wealth in the model.

[16] Figure 2.A.3 plots the profiles for a longer life-cycle horizon and also for consumption and earnings.

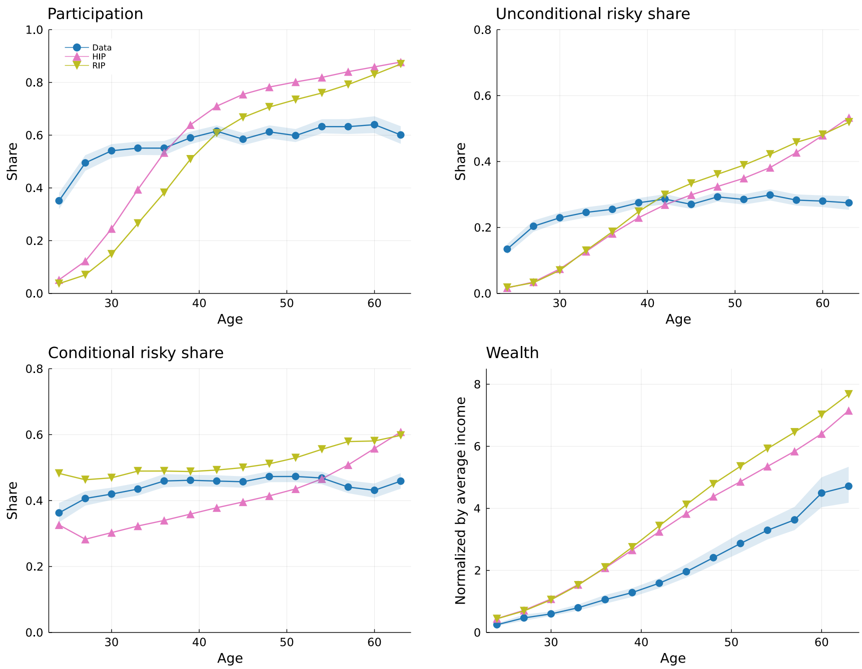


Figure 2.2: Average simulated life-cycle profiles.

As expected, the shapes of the profiles for both RIP and HIP are overall similar to those found by Catherine (2021). Participation is lower at young ages and slightly higher after forty years old compared to the data, but it has the same shape. The same holds for the unconditional risky share. The conditional risky share, instead, matches quite well both the level and the shape of its empirical counterpart. Wealth is higher than in the data across the whole life cycle.[17]

Analysing the differences between HIP and RIP, it is clear from the picture that there are no significant discrepancies between the two specifications when looking at the average life-cycle profile of wealth and unconditional risky share. Instead, participation is higher and the conditional risky share lower in the HIP model compared to the RIP case

[17]Given the known difficulties of these models to replicate the empirical patterns and the fact that in this version I do not estimate preference parameters, the match with the data is relatively good.

across the whole working life.

The reason why the wealth (and consequently consumption, as explained better in the next subsection) profiles are similar between HIP and RIP is related to the result that the estimated autocorrelation coefficients ρ of the AR(1) process governing the persistent component of income are very similar in both models, as previously described in Section 2.3. As shown by Carroll (1992) and Gourinchas and Parker (2002), consumption growth parallels income growth in a life-cycle model where income shocks are very persistent, which translates into very similar wealth profiles.

Turning to the conditional risky share, it is lower in the HIP model because of the heterogeneity in income growth rates implied by the different β_i, which determine the slope of the age trend in the individual-specific part of the income process. Specifically, agents in the HIP model with higher than average β_i - recall that the βs are known from the beginning of life and there is no uncertainty on them - while having higher expected income in levels, because of cyclical skewness they also have a larger part of their total wealth (defined as financial wealth plus human capital) that is risky. Thus, as shown in Catherine (2021), to hedge against this, they optimally choose a lower risky share. In turn, because these agents participate more in the market for the risky asset, this results in a lower average conditional risky share.

Regarding participation, the reason why it is higher in the HIP model has to do again with the heterogeneity in income growth rates. More in detail, imagine there is a wealth-to-income level above which it is optimal for agents to enter the market of the risky asset. If - as in the HIP model - there is a fraction of agents who have lower income growth than the average, following the same argument in the previous paragraph, because they have both lower expected income but also a smaller risky part of total wealth, they have a lower wealth-to-income threshold, and they thus participate more in the market for the risky asset. In turn, because the opposite mechanism for people with higher than average income growth is not as strong (as they accumulate wealth faster), these individuals push upwards the average participation rate.

Although the evidence presented in this section has shown that HIP and RIP imply different profiles of average participation and conditional risky share over the life cycle, just looking at these two schedules is not sufficient to tell them apart. Indeed, while the RIP model matches slightly better both the level and the slope of the conditional risky share, it also overshoots it more at older ages and it also undershoots more participation at the beginning of the life cycle.

The next two sections will thus investigate more in detail the reasons behind the similarities and differences in the profiles generated by the HIP and RIP models just described.

2.4.2 Consumption mean and variance over the life cycle

Even though life-cycle moments of consumption are not usually the interest of portfolio choice models, since previous literature has used the life-cycle patterns of the cross-sectional mean and variance of consumption to discern between HIP and RIP (Guvenen, 2007), in this Section I look more in detail at the model response of these two quantities.

Figure 2.3 plots these profiles. Average log consumption features the usual hump-shaped pattern due to consumption smoothing and the variance is increasing over the life cycle.[18]

[18] Figure 2.A.4 also compares them against consumption data obtained from Krueger and Perri (2006). Despite the fact that the levels of the model-generated profiles do not match the data - the estimation matched the life-cycle variance of earnings from a different dataset featuring higher levels than those in their paper - the shapes are right. More details on the data and on the computation of the empirical profiles can be found in Appendix 2.B.

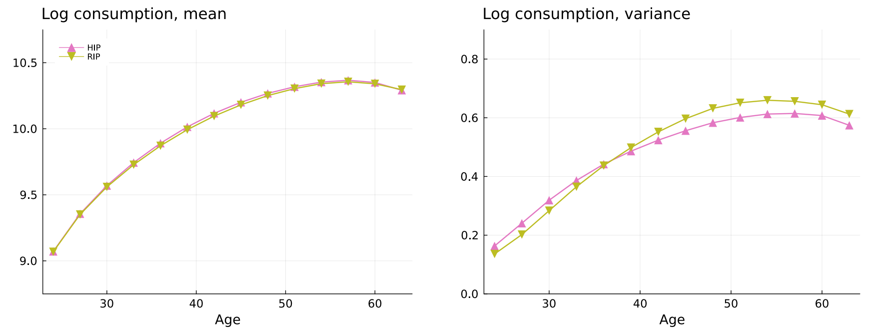


Figure 2.3: Life-cycle profiles of cross-sectional mean and variance of log consumption.

Importantly, the graph confirms the similarity of the two profiles between HIP and RIP previously found for wealth. As previously explained, when the persistence of income shocks is high, in these kinds of models consumption growth closely tracks income growth. Because the effect on average consumption coming from agents with higher than average income growth is compensated by the opposite one coming from agents with lower than average income growth, what drives the patterns in both specifications are average income growth individuals which, in turn, results in the very similar profiles depicted in the picture.

To sum up, when cyclical skewness is properly taken care of, the estimates of the persistence of the income shocks are very similar between HIP and RIP which, in turn, translates into very similar profiles for the mean and variance over the life-cycle of consumption. Consequently, the identification power to discern HIP and RIP coming from these two series is limited.

2.4.3 Identifying restrictions from portfolio choices

As already described in Section 2.4.1, the average life-cycle profiles of participation and conditional risky share differ between HIP and RIP because of the diverse income growth rates distributions implied by the two models. In this section, I explain more in detail the sources of the differences and present identifying restrictions that can be used to test the two in the data.

To disentangle the average effects previously found, Figure 2.4 plots the life-cycle patterns for participation and conditional risky share for HIP and RIP averaging across agents classified according to their average income growth rate during working life.[19][20] The number of the group is increasing in the average growth rate: agents with the lowest growth rates are in group 1. Note that the values depicted for the conditional risky share are expressed in relative terms to group 3.[21]

Considering first participation, the pictures clearly show that the levels for individuals in the high and medium income growth groups are very similar while individuals in the low part of the distribution of income growth rates participate more in the HIP model. The graphs, therefore, confirm the explanation provided in section 2.4.1: because of cyclical skewness, in the HIP model individuals with lower than average income growth have an expected smaller risky component of total wealth, which is balanced by higher participation rates. At young ages, when the human capital component of total wealth is large, this force is so strong that there is almost no heterogeneity across the groups. Conversely, because in both models individuals with high income growth rates accumulate wealth fast - and thus reach faster the relevant wealth-to-income threshold for participation - their rates are not very much different between the two specifications. Another interesting difference is the dispersion of the profiles: agents tend to have more similar profiles across groups in the HIP model.

[19]Figure 2.A.6 plots the profiles for all ages and also for the other variables.

[20]I split the distribution of income growth rates in the two models into five groups: agents below the 20^{th} percentile, between the 20^{th} and the 40^{th} percentiles, between the 40^{th} and the 60^{th} percentiles, between the 60^{th} and the 80^{th} percentiles and above the 80^{th} percentile.

[21]Figure 2.A.6 reports the levels for all variables and a longer life-cycle horizon.

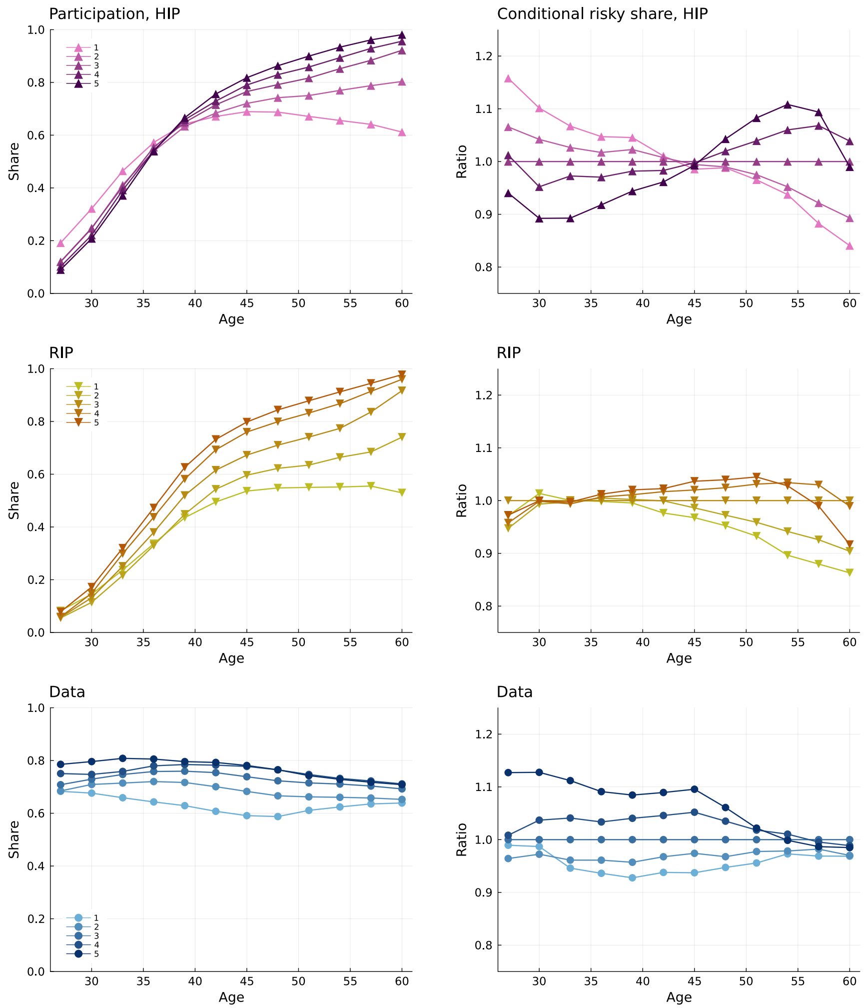


Figure 2.4: Average simulated life-cycle profiles conditioning on average income growth rate over individuals' working lives in HIP and RIP vs. data. Group numbers are increasing in the growth rate, with 1 containing the individuals with growth rates in the lowest part of the income growth rate distribution and 5 in the highest part. Values for conditional risky share are expressed relatively to group 3.

Inspecting the conditional risky share, the graph reveal instead a "butterfly pattern" in the HIP model which is absent in the RIP case. Indeed, individuals in higher income growth groups have lower conditional risky shares than agents in lower groups until around age 40, when this pattern is reversed. The picture, therefore, supports the explanation provided when looking at the average life-cycle profiles: in the HIP model, because of cyclical skewness, individuals with higher than average income growth have an expected higher risky part of total wealth, which is hedged with lower risky share. The opposite holds for agents in low income growth groups. Again, this mechanism is stronger at young ages, when the human capital component of total wealth is large. As individuals age and accumulate wealth, however, they self-insure and, together with the fact that the share of human capital in total wealth gets smaller, they take more risk, which explains the reversion of the pattern at around age 40.

Summing up, this section has shown that, while - as explained in Section 2.4.2 - there does not seem to be much identification power to test HIP and RIP using consumption moments over the life-cycle, testable identifying restrictions can be found by looking at portfolio choices of participation and conditional risky share over the distribution of life-cycle average income growth rates. Testing these restrictions in the data is exactly what I do in the next section.

2.4.4 Testing the restrictions in the data

The model-based implications of the two specifications depicted in Figure 2.4 and described above can be tested empirically with panel data on individual income and wealth. Because the Survey of Consumer Finances is constructed as a repeated cross-section, it cannot be directly used for this purpose. Thus, the content of this section relies on data from the Swedish Wealth and Income Registry spanning the period 1994-2015.[22][23] These data include yearly variables on demographic

[22] As the model is estimated on US data, the underlying assumption behind this comparison is that the patterns in the Swedish context are not very different from those in the US.

[23] For more details on this comprehensive dataset, see Catherine et al. (2021) and the references cited therein.

characteristics, income, and wealth holdings at the individual level for the whole universe of Swedish residents.[24] For the purpose of my analysis, the variables needed are age, a measure of income, and a measure of the risky share, all at the individual level. Regarding income, the variable used is a series of non-financial disposable income based on the definition by Statistics Sweden, which spans the period 1994-2015. Following Catherine et al. (2021), instead, the risky share is defined as the ratio between the sum of wealth invested in stocks and funds over the sum of these two variables with cash.[25][26]

Given the large number of individuals in the simulation and the fact that agents know from the beginning of the life cycle the parameters governing the deterministic part of their income, classifying them into different average income growth groups can be done straightforwardly by computing the average yearly income log growth rate over their working life. The same is not true in the data because of the presence of other confounding factors (e.g., agents' expectations, cohort and year effects, etc.). Therefore, the following procedure is adopted. First, the average log growth of income in all the working age years (from 23 to 65 years old) is computed for each individual, using the whole sample of available income data, i.e., 1994-2015. Second, this measure is regressed on dummy variables controlling for the agent's age in the first year of available wealth data. The resulting residual is then used as the relevant measure for classifying individuals in different average income growth groups. Finally, the life-cycle profiles for participation and conditional risky share are computed using the procedure described in

[24]The wealth data used cover the years 2000-2007. They include bank accounts, mutual funds, and holdings of stocks, bonds, and derivatives and they were collected due to a wealth tax. A detailed description of this dataset is available in Bach et al. (2020).

[25]Financial wealth and its components are defined as in Bach et al. (2020).

[26]The final sample used includes individuals between 23 and 91 years old, without business income, with income and financial wealth above 10000 SEK (in 2015 terms) and below the 0.999 percentile of their respective distributions. Furthermore, only individuals without large real estate transactions and present in all the years in which income and wealth data, respectively, are available are considered. Finally, individuals with income-to-financial wealth ratios below and above the 0.15 and 0.85 percentiles of the distribution (respectively, about 0.5 and 10) of this variable are also excluded.

Appendix 2.B for all the five groups.[27]

The bottom part of Figure 2.4 reports the results.[28] Participation is relatively stable across the life cycle, and, in general, slightly increasing at young ages. The pattern emerging across the different income growth groups is clear: participation for group 1 is always the lowest and the other groups follow in increasing order. A similar trend (except for group 1 at the very beginning and for above 50 years old, where the lines start to overlap) is also visible when considering the conditional risky share.

What can be inferred from these empirical moments? Despite the fact that the level in both model specifications at the beginning of the life cycle is lower than in the data[29], the order and variation of the participation rate between groups at young ages in the data resemble more the RIP case. Analysing the conditional risky share, instead, reveals that the data clearly do not support the "butterfly pattern" found for the HIP specification. Even though the RIP case shows a counterfactual overlap between the groups at young ages, in this specification an order and variation of the schedules among groups more in line with the data is achieved earlier in the life cycle.

Summing up, although testing the restrictions in the data in the way described in this section has not delivered a conclusive answer, the data seem to support slightly more the RIP hypothesis, especially because no evidence of a "butterfly pattern" for the conditional risky share was found.

---

[27]I kindly thank Paolo Sodini for sharing these moments.

[28]The levels for the conditional risky share are reported in Figure 2.A.7. For both participation and the conditional risky share the values obtained are in line with what reported in Catherine et al. (2021).

[29]This is even more clear with Swedish data where, compared to the pattern obtained with data from the Survey of Consumer Finances previously described, participation is higher at young ages. The level of the conditional risky share, instead, is quite similar.

2.5 Conclusion

In this paper I have investigated what inference on the properties of the income process can be drawn from a state of the art portfolio choice model.

First, I have documented that cyclical skewness needs to be included in the stochastic process for income in order to correctly estimate the amount of risk deriving from the persistence of the shocks. Indeed, the HIP model without cyclical skewness overestimates the share of risk attributed to the heterogeneity in life-cycle profiles and underestimates the share deriving from persistence.

Second, when the income process includes cyclical skewness, I find that the estimated autocorrelation coefficient in the AR(1) process for the persistent component of income is similar for HIP and RIP. Because consumption growth parallels income growth in life-cycle models when the persistence of the shocks is high this, in turn, implies very similar consumption and wealth profiles in both specifications. Therefore, the cross-sectional mean and variance of the wealth and consumption profiles over the life-cycle do not have enough identification power to disentangle between HIP and RIP.

Third, I have documented that the patterns of participation and conditional risky share across the distribution of average working life income growth rates have identification power to discern between HIP and RIP. Specifically, compared to the RIP case, the distribution of income growth rates in the HIP process determines less heterogeneity across income groups for participation and a "butterfly pattern" for the conditional risky share, which can both be tested empirically. Although the data did not deliver a conclusive answer, more support was found for the RIP case, especially because no evidence of the "butterfly pattern" was discoverd.

This work opens the avenue to future research in several ways. While this paper has used it to reach other conclusions, the result that the persistence of income shocks is similar across HIP and RIP when cyclical skewness is taken care of is very interesting and I am investigating it more in detail in ongoing research. Furthermore, I have focused only on a specific set of moments: additional work is needed to check whether other moments contain useful information for identification. In addition, the data patterns used for testing the restrictions considered the average agent in each income growth group: it would be interesting to test the robustness of the results for agents differently exposed to cyclical skewness (Catherine et al., 2021), or with different wealth-to-income ratios. Lastly, this study has focused on a particular question, namely inferring from portfolio choices whether the income process is more in line with the HIP or RIP hypothesis. Future work could use the approach outlined in this paper to look at income properties more generally as revealed by portfolio choices.

References

Arellano, M., Blundell, R., and Bonhomme, S. (2017). Earnings and consumption dynamics: A nonlinear panel data framework. *Econometrica*, 85(3):693–734.

Arellano, M., Bonhomme, S., De Vera, M., Hospido, L., and Wei, S. (2022). Income risk inequality: Evidence from spanish administrative records. *Quantitative Economics*, 13(4):1747–1801.

Bach, L., Calvet, L. E., and Sodini, P. (2020). Rich pickings? risk, return, and skill in household wealth. *American Economic Review*, 110(9):2703–47.

Benzoni, L., Collin-Dufresne, P., and Goldstein, R. S. (2007). Portfolio choice over the life-cycle when the stock and labor markets are cointegrated. *The Journal of Finance*, 62(5):2123–2167.

Blundell, R., Pistaferri, L., and Preston, I. (2008). Consumption inequality and partial insurance. *American Economic Review*, 98(5):1887–1921.

Blundell, R. and Preston, I. (1998). Consumption inequality and income uncertainty. *The Quarterly Journal of Economics*, 113(2):603–640.

Busch, C., Domeij, D., Guvenen, F., and Madera, R. (2022). Skewed idiosyncratic income risk over the business cycle: Sources and insurance. *American Economic Journal: Macroeconomics*, 14(2):207–42.

Calvet, L. E., Campbell, J. Y., Gomes, F. J., and Sodini, P. (2021). The cross-section of household preferences. *Working Paper*.

Carroll, C. D. (1992). The Buffer-Stock Theory of Saving: Some Macroeconomic Evidence. *Brookings Papers on Economic Activity*, 23(2):61–156.

Catherine, S. (2021). Countercyclical Labor Income Risk and Portfolio Choices over the Life Cycle. *The Review of Financial Studies*. hhab136.

Catherine, S., Sodini, P., and Zhang, Y. (2021). Countercyclical income risk and portfolio choices: Evidence from sweden. *Working Paper*.

Chang, Y., Hong, J., Karabarbounis, M., and Wang, Y. (2018a). Income Volatility and Portfolio Choices. 2018 Meeting Papers 412, Society for Economic Dynamics.

Chang, Y., Hong, J. H., and Karabarbounis, M. (2018b). Labor market uncertainty and portfolio choice puzzles. *American Economic Journal: Macroeconomics*, 10(2):222–62.

Cocco, J. F., Gomes, F. J., and Maenhout, P. J. (2005). Consumption and Portfolio Choice over the Life Cycle. *The Review of Financial Studies*, 18(2):491–533.

De Nardi, M., Fella, G., and Paz-Pardo, G. (2019). Nonlinear Household Earnings Dynamics, Self-Insurance, and Welfare. *Journal of the European Economic Association*, 18(2):890–926.

Deaton, A. and Paxson, C. (1994). Intertemporal choice and inequality. *Journal of Political Economy*, 102(3):437–467.

Fagereng, A., Guiso, L., and Pistaferri, L. (2017). Portfolio Choices, Firm Shocks, and Uninsurable Wage Risk. *The Review of Economic Studies*, 85(1):437–474.

Gomes, F. and Michaelides, A. (2005). Optimal life-cycle asset allocation: Understanding the empirical evidence. *The Journal of Finance*, 60(2):869–904.

Gomes, F. and Michaelides, A. (2008). Asset Pricing with Limited Risk Sharing and Heterogeneous Agents. *The Review of Financial Studies*, 21(1):415–448.

Gourinchas, P.-O. and Parker, J. A. (2002). Consumption over the life cycle. *Econometrica*, 70(1):47–89.

Guvenen, F. (2007). Learning your earning: Are labor income shocks really very persistent? *The American Economic Review*, 97(3):687–712.

Guvenen, F. (2009). An Empirical Investigation of Labor Income Processes. *Review of Economic Dynamics*, 12(1):58–79.

Guvenen, F., Karahan, F., Ozkan, S., and Song, J. (2021). What do data on millions of u.s. workers reveal about lifecycle earnings dynamics? *Econometrica*, 89(5):2303–2339.

Guvenen, F., Ozkan, S., and Song, J. (2014). The nature of countercyclical income risk. *Journal of Political Economy*, 122(3):621–660.

Guvenen, F. and Smith, A. A. (2014). Inferring labor income risk and partial insurance from economic choices. *Econometrica*, 82(6):2085–2129.

Hall, R. E. and Mishkin, F. S. (1982). The sensitivity of consumption to transitory income: Estimates from panel data on households. *Econometrica*, 50(2):461–481.

Heathcote, J., Storesletten, K., and Violante, G. L. (2005). Two views of inequality over the life cycle. *Journal of the European Economic Association*, 3(2/3):765–775.

Huggett, M. and Kaplan, G. (2016). How large is the stock component of human capital? *Review of Economic Dynamics*, 22:21 – 51.

Inkmann, J., Lopes, P., and Michaelides, A. (2010). How Deep Is the Annuity Market Participation Puzzle? *The Review of Financial Studies*, 24(1):279–319.

Kaplan, G. and Violante, G. L. (2010). How much consumption insurance beyond self-insurance? *American Economic Journal: Macroeconomics*, 2(4):53–87.

Krueger, D. and Perri, F. (2006). Does Income Inequality Lead to Consumption Inequality? Evidence and Theory. *The Review of Economic Studies*, 73(1):163–193.

Lillard, L. A. and Willis, R. J. (1978). Dynamic aspects of earning mobility. *Econometrica*, 46(5):985–1012.

MaCurdy, T. E. (1982). The use of time series processes to model the error structure of earnings in a longitudinal data analysis. *Journal of Econometrics*, 18(1):83–114.

Meghir, C. and Pistaferri, L. (2004). Income variance dynamics and heterogeneity. *Econometrica*, 72(1):1–32.

Merton, R. (1969). Lifetime portfolio selection under uncertainty: The continuous-time case. *The Review of Economics and Statistics*, 51:247–57.

Storesletten, K., Telmer, C., and Yaron, A. (2004). Cyclical dynamics in idiosyncratic labor market risk. *Journal of Political Economy*, 112(3):695–717.

Storesletten, K., Telmer, C. I., and Yaron, A. (2007). Asset pricing with idiosyncratic risk and overlapping generations. *Review of Economic Dynamics*, 10(4):519–548.

Appendices

2.A Additional figures

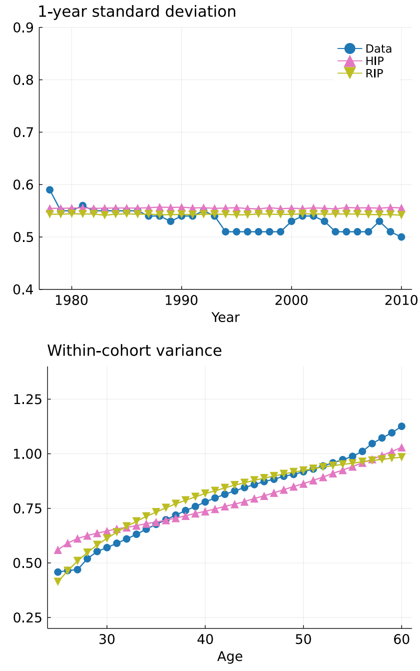


Figure 2.A.1: Individual income moments, comparison between different model specifications without cyclical skewness and data.

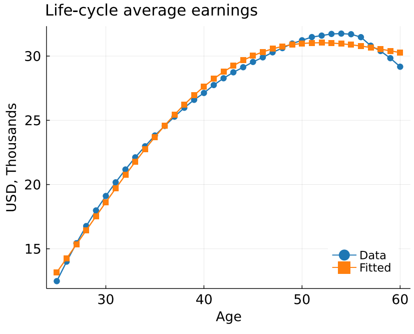


Figure 2.A.2: Fitted age polynomial vs. data

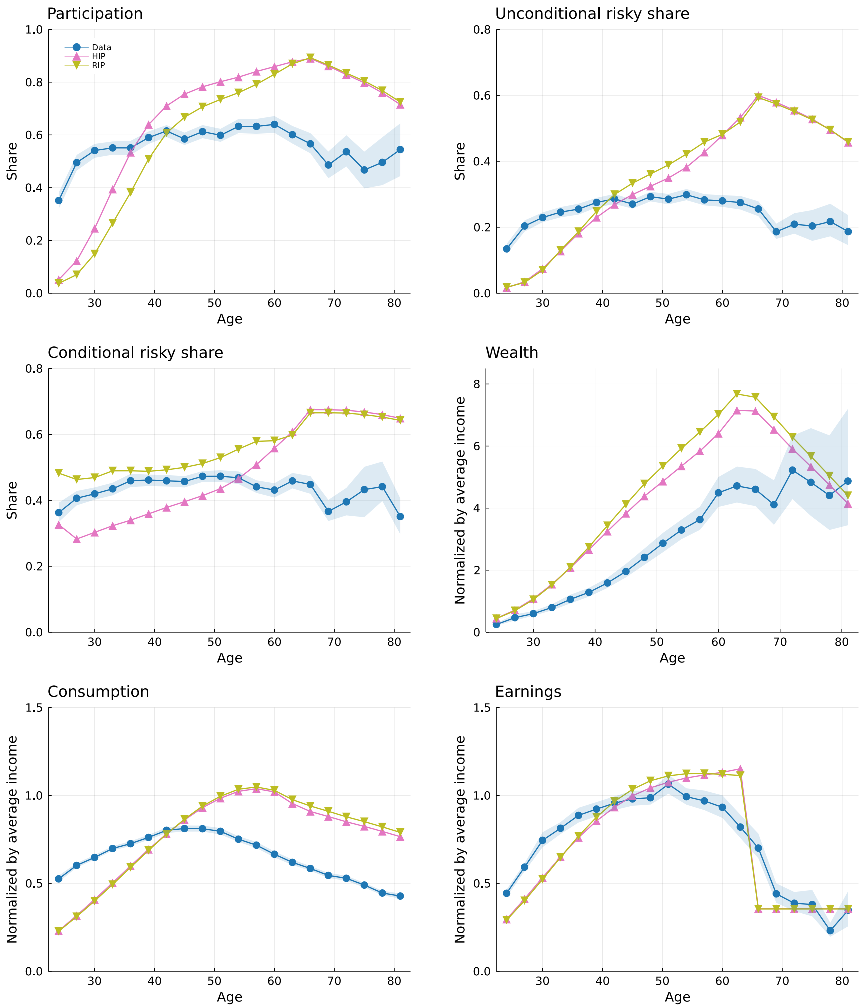


Figure 2.A.3: Average simulated life-cycle profiles from the model, all ages and variables. Consumption data are from Krueger and Perri (2006) and data for the other variables from the Survey of Consumer Finances. More details on the data can be found in Appendix 2.B.

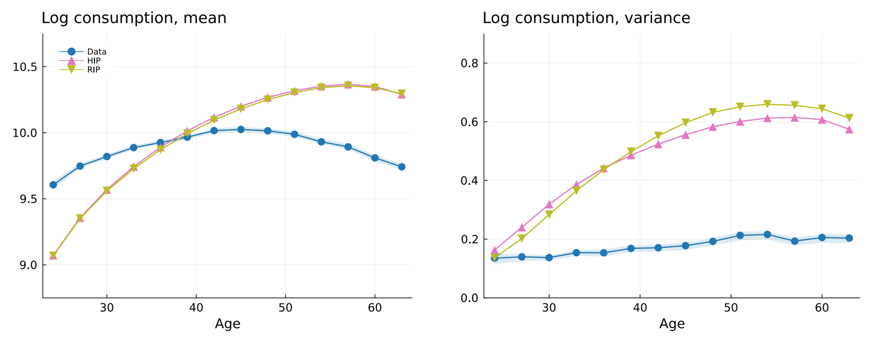
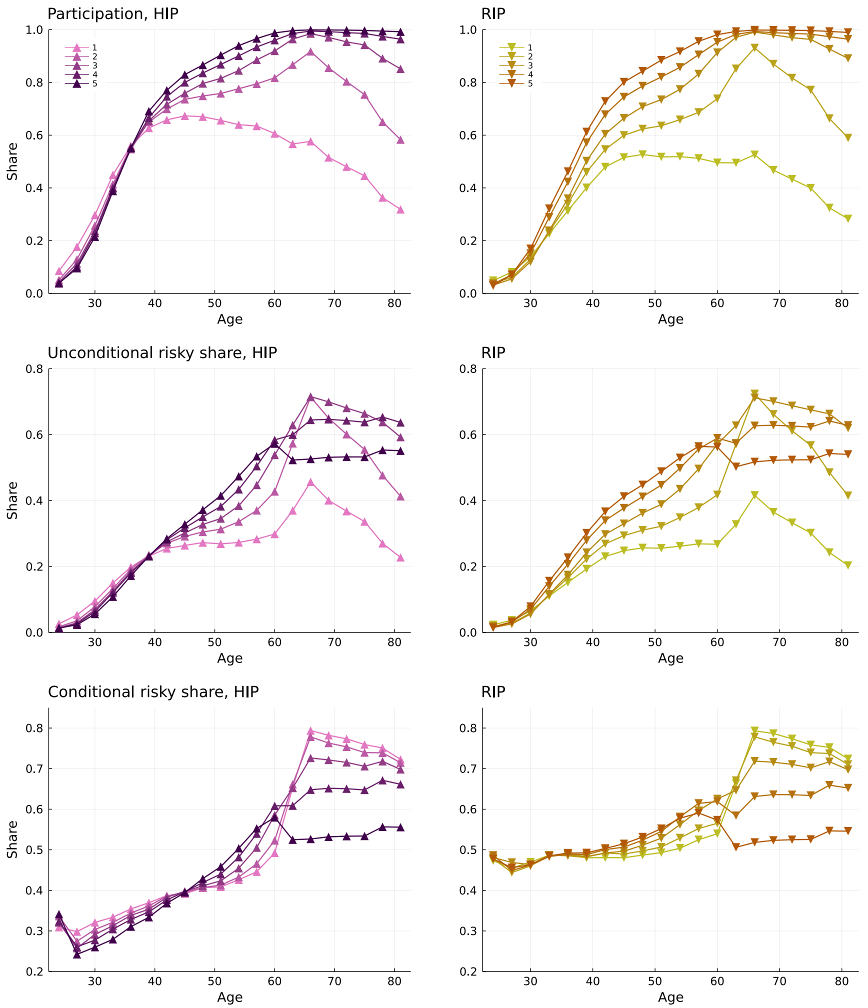


Figure 2.A.4: Life-cycle profiles of cross-sectional mean and variance of log consumption vs. data from Krueger and Perri (2006).


Participation, HIP
RIP
Unconditional risky share, HIP
RIP
Conditional risky share, HIP
RIP
Share
Age
1
2
3
4
5

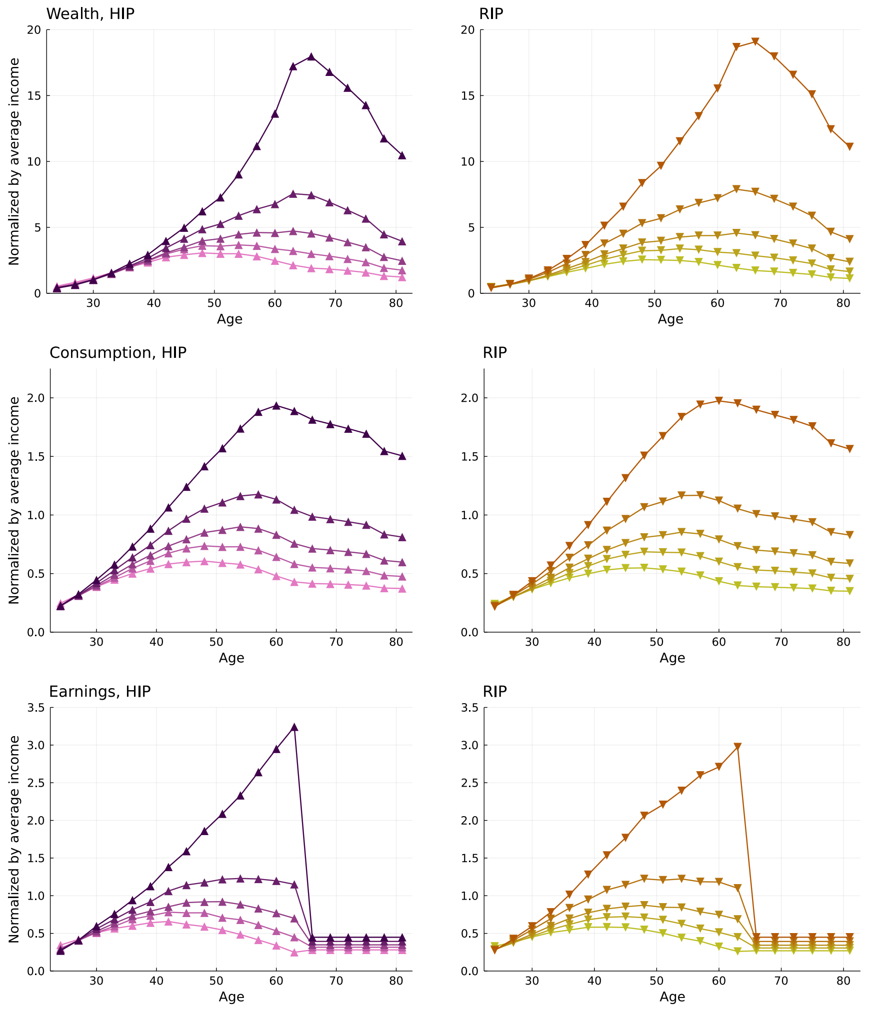


Figure 2.A.6: Average simulated life-cycle profiles conditioning on average income growth rate over individuals' working lives, all ages and variables, HIP and RIP. Group numbers are increasing in the growth rate, with 1 containing the individuals with growth rates in the lowest part of the income growth rate distribution and 5 in the highest part.

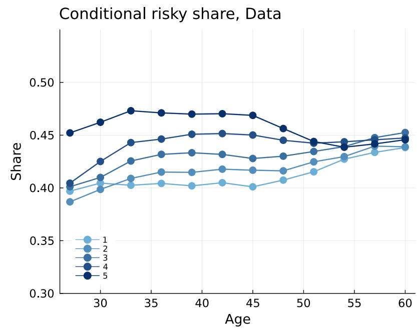


Figure 2.A.7: Average life-cycle profile for conditional risky share conditioning on average income growth rate over individuals' working lives, data. Group numbers are increasing in the growth rate, with 1 containing the individuals with growth rates in the lowest part of the income growth rate distribution and 5 in the highest part.

2.B Data

This section describes in detail the data sources used in the paper.

Demographics. Survival probabilities are taken from the 2019 actuarial life table compiled by US Social Security Administration, available at this link `https://www.ssa.gov/oact/STATS/table4c6.html`.

Aggregate variables. Data for the return on the risky asset are taken from Robert Shiller's website (`http://www.econ.yale.edu/~shiller/data/ie_data.xls`). I use the variable "Real Total Return Price". I get a yearly series by taking the monthly value at the beginning of the year. The log return in year t is then the log difference between the yearly price at $t+1$ and t. For the wage index I use two sources. From 1944 to 1978 I compute log growth rates using data from Emmanuel Saez available at this link `http://eml.berkeley.edu/~saez/TabFig2012prel.xls`. Specifically, I use the variable "Average wage income ($ latest year)" in Table B1. For the period 1979-2011 I use the log growth rates reported by Guvenen et al. (2014) in Table A1 (available at this link `https://fatihguvenen.com/s/gos-jpe2014-data.xlsx`) for the variable "Change in log earnings averaged over workers x100". To deflate nominal variables I use the CPI index for the US (reference year 2010) from the World Bank, which is available at their online database.

Individual income estimation. For estimation of the stochastic process governing individual income I use again the values reported by Guvenen et al. (2014) (available at this link `https://fatihguvenen.com/s/gos-jpe2014-data.xlsx`). Specifically, I use the time series for standard deviation of earnings growth at the 1 and 5 year horizons reported in Table A8 and the time series for the 10th, 50th and 90th percentiles of the distribution of earnings growth at the 1, 3 and 5 year horizons reported in Table C1. In addition, I use within-cohort variances of log earnings and average log earning by age reported, respectively, in the sheets "Figure D3" and "Figure C36" compiled by Guvenen et al. (2021), available at this link `https://fatihguvenen.com/s/gkos_2021_moments.xlsx`.

Agents' balance sheets. For variables related to agents' balance sheets I rely on the eleven waves of the Survey of Consumer Finances from 1989 to 2019. More in detail, I use the "summary extract public data", which are available at the Federal Reserve's website. In the description below, variables in italics refer to variable names in the original datasets. Additionally, to ensure comparability across different surveys, I do not use the absolute weights provided in the original data, but their rescaled version (i.e., the original weights divided by their sum in each year). The variables I focus on are: labor earnings (*wageinc*), net worth (*networth*)[30], financial wealth (*fin*), equity (*equity*)[31].

[30]This is the sum of financial assets (cash, savings, retirement, investment accounts, etc.), businesses and residential assets, minus all debts.

[31]This is the sum of directly held stocks and stocks own indirectly through mutual funds and retirement accounts. The survey asks households whether these accounts are invested mostly into bonds or stocks and imputes a fraction of the total value of the account to the equity variable based on the response.

I restrict the sample to households between age 23 and 82 and, in order to filter out entrepreneurs and self-employed people, I remove all the households for whom the variable *bus* is not zero. I further filter out individuals whose labor earnings are lower than 1000 USD and whose net worth is lower than 1000 USD.[32]

Additionally, I create the following variables:

- the ratio between financial wealth and average income, where the latter is defined as the yearly cross-sectional average labor income (for comparability with the model's variables) computed using survey weights, that is $fin/\mathbb{E}\left[wageinc\right]$;
- the ratio between income minus a 15% tax and average income, that is $wageinc/\mathbb{E}\left[wageinc\right]$;
- risky share defined as the ratio between stock holdings and financial wealth, that is $equity/fin$;
- participation defined as a dummy equal to one if the risky share is strictly positive.

Agents' consumption. I use the data compiled by Krueger and Perri (2006).[33] In the description below, variables in italics refer to variable names in their dataset. In addition to the sample restrictions already present in the dataset available for download (incomplete income respondents, households who report 0 USD in food consumption, households who only report only food consumption), I apply other similar restrictions as they do in their paper: I remove observations with positive labor income but no hours worked and I restrict to households completing all the interviews. Then, as they do in their paper, I classify an household as belonging to year t if the last interview was conducted

[32]The Survey of Consumer Finances' extracts at the time of download are in 2019 USD.

[33]The data are available at Fabrizio Perri's website `http://www.fperri.net/research_data.htm`.

between the second quarter of year t and the first quarter of year $t + 1$, I define yearly income values as those reported in the last interview and and yearly consumption as the sum of quarterly consumption reported in each of the four interviews. The income measure I use is total income before taxes (*incbetax*) and the relevand consumption measure I consider is the one constructed by the authors (*ndpbe0*). In addition, I also define yearly survey weights for each observation as the sum of survey weights for each of the four interviews. Furthermore, age is defined as the age of the reference person (*refage*) at the time of the last interview. I restrict the sample to households with at least 1000 USD of wealth (defined as the sum of financial wealth (*finwea*) and the value of owned residence (*propval*), at least 1000 USD of income, for which both the reference person and the spouse have zero business income (*refby* and *spoby* equal to zero) and for which age is between 23 and 82.

In addition to the variables already provided in the dataset, I construct an additional variable: the ratio between consumption and average income, where the latter is defined as the yearly cross-sectional average total income before taxes (for comparability with the model's variables) computed using survey weights, that is *ndpbe0*$/\mathbb{E}\left[\textit{incbetax}\right]$.

Life-cycle profiles. To construct life-cycle profiles I use a method similar to Heathcote et al. (2005). First, I build 3-year age groups. Then, I compute $m_{a,c,t}$, that is, moment m for households in age group a, with cohort c in year t using survey weights. I then regress these moments on age group and year dummies[34] and recover the age profile for moment m by adding the unconditional average of the coefficients of the time dummies to the coefficients of the age group dummies. Standard errors on such moments are computed by boostrapping the data 1000 times at the cohort and year level.

[34]This is what Heathcote et al. (2005) call the "time view".

2.C Numerical solution

Discretization and grids construction.

Normally distributed random variables. Let X be an i.i.d. Normally distributed random variable with mean μ_x and standard deviation σ_x. I discretize X using Gaussian quadrature. Specifically, the support of X is approximated with a finite grid of values $x_1, \ldots, x_{N_q}$ computed as follows:

$$x_j = \mu_x + \sqrt{2}\sigma_x Z_j, \quad j = 1, \ldots, N_q$$

where the Z_j's are Gauss-Hermite nodes and the probability mass of each point of the discretized support is computed as:

$$p(x_j) = \omega_j / \sqrt{\pi}, \quad j = 1, \ldots, N_q$$

where the ω_j's are Gauss-Hermite weights. This procedure is used to discretize $r_{2,t}$, ϕ_t, and the distributions of $\nu_{i,t}$ and $r_{1,t}$ conditional, respectively, on tail/non-tail event and on stock market crash/normal period. N_q is the same for all shocks.

Persistent component of idiosyncratic income. I approximate the process governing the evolution of $z_{i,t}$ as follows: (i) I discretize the conditional distribution of $\varepsilon_{i,t}$, (ii) I compute the evolution of the persistent component according to equation (2.8) and (iii) I evaluate the model functions at the resulting value of $z_{i,t}$. The advantage of this method is that it requires to discretize just the conditional distribution of $\varepsilon_{i,t}$, which is easier than discretizing the full process of $z_{i,t}$. In particular, the crucial connections between the higher moments of $z_{i,t}$ and other variables are preserved. The disadvantage is that the resulting values of $z_{i,t}$ will very often be off grid, so I need a grid of values that captures well the behavior of the model at such points given the interpolation procedure.[35] Given the above discussion, the grids for $\varepsilon_{i,t}$ and $z_{i,t}$ are constructed as follows. The conditional distribution of $\varepsilon_{i,t}$ is discretized using the procedure described above for Normally distributed shocks with N_q points. To set up the grid for $z_{i,t}$, instead, I first construct an exponentially spaced grid of $(N_z - 1)/2 + 1$ points with minimum value equal to zero, maximum value equal to $z_{\max}$ and spacing parameter equal to spacing_z. This gives me the positive side of the grid plus the central point (which is therefore equal to zero). Then, I add the negative $(N_z - 1)/2$ values by taking the negative of the positive values just computed and obtain the full grid of N_z points.

Average income. The grid for average income $\bar{L}_{i,t}$ is an exponentially spaced grid of $N_{\bar{L}}$ points with minimum value equal to the lowest possible realization of $\bar{L}$ implied by the income process and the formula for average income in (2.16) and maximum value equal to the highest possible realization of $\bar{L}$ implied by the formula, that is 2.5. The spacing parameter is equal to $\text{spacing}_{\bar{L}}$.

Life-cycle parameters. The grid for α_i is an exponentially spaced grid of $(N_\alpha - 1)/2 + 1$ points with minimum value equal to zero, maximum value equal to $\alpha_{\max}$ and spacing parameter equal to $\text{spacing}_{\text{lc}}$. This gives me the positive side of the grid plus the central point (which is therefore equal to zero). Then, I add the negative $(N_\alpha - 1)/2$ values by taking the negative of the positive values just computed and obtain the full grid of N_α points. The grid for β_i is constructed with the same procedure, using N_β points, maximum value $\beta_{\max}$ and spacing parameter $\text{spacing}_{\text{lc}}$.

Cash on hand and savings. The grid for cash on hand is an exponentially spaced grid of $N_{\hat{a}}$ points with minimum value equal to the lowest possible realization of cash on hand implied by the model, maximum value equal to $\hat{a}_{\max}$ and spacing parameter equal to $\text{spacing}_{\hat{a}}$. The

[35] See below for more details on the interpolation method.

grid for savings $\hat{s}_{i,t}$ is an exponentially spaced grid of $N_{\hat{s}}$ points with minimum value equal to $\overline{s}$, maximum value equal to $\hat{s}_{\max}$ and spacing parameter equal to $\text{spacing}_{\hat{s}}$.[36]
Table 2.C.1 summarizes the choices for the numerical parameters.

Parameter	Value	Description
Panel A: numerical parameters for model solution		
N_q	3	Number of points Gaussian quadrature
N_z	15	Number of points grid persistent/idiosyncratic income
$N_{\hat{a}}$	51	Number of points grid cash on hand
$N_{\hat{s}}$	$N_{\hat{a}}$	Number of points grid savings
$N_{\bar{L}}$	21	Number of points grid average income
N_α	5	Number of points grid life-cycle constant
N_β	5	Number of points grid life-cycle slope
$z_{\max}$	4.5	Maximum value grid persistent income
$\hat{a}_{\max}$	200.0	Maximum value grid cash on hand
$\hat{s}_{\max}$	$\hat{a}_{\max}$	Maximum value grid savings
$\alpha_{\max}$	$3\sigma_\alpha$ in HIP model	Maximum value grid life-cycle constant
$\beta_{\max}$	$3\sigma_\beta$ in HIP model	Maximum value grid life-cycle slope
spacing_z	1.5	Spacing parameter grid persistent income
$\text{spacing}_{\hat{a}}$	1.25	Spacing parameter grid cash on hand
$\text{spacing}_{\hat{s}}$	$\text{spacing}_{\hat{a}}$	Spacing parameter grid savings
$\text{spacing}_{\bar{L}}$	1.25	Spacing parameter grid average income
$\text{spacing}_{\text{lc}}$	1.25	Spacing parameter grids life-cycle parameters
Panel B: numerical parameters for model simulation		
T_{eco}	1000	Number of different time-series of aggregate shocks to simulate
N_{sim}	1500	Number of agents to simulate
Panel C: numerical parameters for estimation		
N_{glo}	1000	Number of points to evaluate in global stage (5000 for agg. proc.)
N_{loc}	10	Number of points to evaluate in local stage (50 for agg. proc.)
N_{eco}	5	Number of economies to simulate
T_{cal}	10^5	Number of time-series points to simulate for aggregate shocks
T_{dis}	1000	Number of periods to discard for moments computation
N_{cal}	1500	Number of individuals in each cohort to simulate

Table 2.C.1: Numerical parameters.

Solving the optimization problem. Whenever it does not lead to confusion I am dropping α_i, β_i and $\bar{L}_{i,t}$ from the state variable. Also, for ease of exposition, I will consider the case in which disposable

[36] Variables with a hat on top refer to normalized variables as defined in section 2.C below.

income coincides with labor income and there are no participation costs: the general case is a straightforward extension. Because the aggregate component of the wage follows a random walk, it is possible to rescale the problem to get rid of w as a state variable as follows. Let $\hat{x}_{i,t} = x_{i,t}/e^{w_t+\bar{f}(t)}$ for a generic variable[37] x and $\hat{V}_{i,t}\left(a_{i,t}, z_{i,t}\right) := V_{i,t}\left(a_{i,t}, z_{i,t}, 0\right)$, so that I can write:

$$V_{i,t}\left(a_{i,t}, z_{i,t}, w_t\right) = e^{w_t+\bar{f}(t)} V_{i,t}\left(\frac{a_{i,t}}{e^{w_t+\bar{f}(t)}}, z_{i,t}, 0\right) = e^{w_t+\bar{f}(t)} \hat{V}_{i,t}\left(\hat{a}_{i,t}, z_{i,t}\right)$$

Using the above definitions, letting $\Delta w_t := w_t - w_{t-1}$ and $\Delta f_t := \bar{f}(t) - \bar{f}(t-1)$, the optimization problem during working life can be rewritten as follows:

$$\hat{V}_{i,t}(\hat{a}_{i,t}, z_{i,t}) = \max_{\xi_{i,t}, \hat{c}_{i,t}, \hat{s}_{i,t}} \Bigg\{ (1-\delta)\hat{c}_{i,t}^{1-\psi}$$
$$+\delta \left[\mathbb{E}_t \left[\left(p_t \hat{V}_{i,t+1}^{1-\gamma}(\hat{a}_{i,t+1}, z_{i,t+1}) + b(1-p_t)\hat{a}_{i,t+1}^{1-\gamma} \right) e^{(\Delta w_{t+1} + \Delta f_{t+1})(1-\gamma)} \right] \right]^{\frac{1-\psi}{1-\gamma}} \Bigg\}^{\frac{1}{1-\psi}}$$

subject to:

$$\begin{aligned} \hat{c}_{i,t} + \hat{s}_{i,t} &= \hat{a}_{i,t} \\ \hat{a}_{i,t+1} &= \left[R_f + \xi_{i,t} R_{t+1}^e \right] \hat{s}_{i,t} e^{-\Delta w_{t+1} - \Delta f_{t+1}} + e^{z_{i,t+1} + \nu_{i,t+1}} \\ \hat{s}_{i,t} &\geq \bar{s} \end{aligned}$$

For the solution, it is useful to define:

$$\begin{aligned} \tilde{V}_{i,t}(\hat{s}_{i,t}, \xi_{i,t}, z_{i,t}) = \Big[\mathbb{E}_t \Big[\Big(& p_t \hat{V}_{i,t+1}^{1-\gamma}(\hat{a}_{i,t+1}, z_{i,t+1}) \\ & + b(1-p_t)\hat{a}_{i,t+1}^{1-\gamma} \Big) e^{(\Delta w_{t+1} + \Delta f_{t+1})(1-\gamma)} \Big] \Big]^{\frac{1-\psi}{1-\gamma}} \end{aligned}$$

[37] That is, I also rescale the problem by $\bar{f}(t)$.

The first order condition with respect to $\xi_{i,t}$ reads:

$$\frac{\partial \tilde{V}_{i,t}(\tilde{s}_{i,t},\xi_{i,t},z_{i,t})}{\partial \xi_{i,t}} = 0 \iff$$
$$\mathbb{E}_t \Bigg[e^{-\gamma(\Delta w_{t+1}+\Delta f_{t+1})} R^e_{t+1}$$
$$\times \left(p_t \hat{V}^{-\gamma}_{i,t+1}(\hat{a}_{i,t+1},z_{i,t+1}) \frac{\partial \hat{V}_{i,t+1}(\tilde{a}_{i,t+1},z_{i,t+1})}{\partial \hat{a}_{i,t+1}} + b(1-p_t)\hat{a}^{-\gamma}_{i,t+1} \right) \Bigg] = 0$$

The first order condition with respect to $\hat{s}_{i,t}$ reads:

$$(1-\delta)(1-\psi)\hat{c}^{-\psi}_{i,t} = \delta \frac{\partial \tilde{V}_{i,t}(\hat{s}_{i,t},\xi_{i,t},z_{i,t})}{\partial \hat{s}_{i,t}} \iff$$
$$(1-\delta)\hat{c}^{-\psi}_{i,t} = \delta \tilde{V}_{i,t}(\hat{s}_{i,t},\xi_{i,t},z_{i,t})^{\frac{\gamma-\psi}{1-\psi}} \times$$
$$\mathbb{E}_t \Bigg[e^{-\gamma(\Delta w_{t+1}+\Delta f_{t+1})} \left(R_f + \xi_{i,t} R^e_{t+1}\right)$$
$$\times \left(p_t \hat{V}^{-\gamma}_{i,t+1}(\hat{a}_{i,t+1},z_{i,t+1}) \frac{\partial \hat{V}_{i,t+1}(\tilde{a}_{i,t+1},z_{i,t+1})}{\partial \hat{a}_{i,t+1}} + b(1-p_t)\hat{a}^{-\gamma}_{i,t+1} \right) \Bigg]$$

Finally, the envelope condition is:

$$\frac{\partial \hat{V}_{i,t}(\hat{a}_{i,t},z_{i,t})}{\partial \hat{a}_{i,t}} = \frac{\hat{V}_{i,t}(\hat{a}_{i,t},z_{i,t})^{\psi}}{1-\psi} \times$$
$$\left[(1-\delta)(1-\psi)\hat{c}^{-\psi}_{i,t} \frac{d\hat{c}_{i,t}}{d\hat{a}_{i,t}} + \delta \left(\frac{\partial \tilde{V}_{i,t}(\hat{s}_{i,t},\xi_{i,t},z_{i,t})}{\partial \hat{s}_{i,t}} \frac{d\hat{s}_{i,t}}{d\hat{a}_{i,t}} + \frac{\partial \tilde{V}_{i,t}(\hat{s}_{i,t},\xi_{i,t},z_{i,t})}{\partial \xi_{i,t}} \frac{d\xi_{i,t}}{d\hat{a}_{i,t}} \right) \right]$$

Using the above first order conditions and the fact that $\frac{d\hat{c}_{i,t}}{d\hat{a}_{i,t}} + \frac{d\hat{s}_{i,t}}{d\hat{a}_{i,t}} = 1$, the envelope condition reduces to:

$$\frac{\partial \hat{V}_{i,t}(\hat{a}_{i,t},z_{i,t})}{\partial \hat{a}_{i,t}} = \hat{V}_{i,t}(\hat{a}_{i,t},z_{i,t})^{\psi}(1-\delta)\hat{c}^{-\psi}_{i,t}$$

Replacing the envelope condition in the FOCs above I obtain:

$$\mathbb{E}_t \Big[e^{-\gamma(\Delta w_{t+1}+\Delta f_{t+1})} R^e_{t+1} \times \left(p_t (1-\delta) \hat{V}^{\psi-\gamma}_{i,t+1}(\hat{a}_{i,t+1},z_{i,t+1}) \hat{c}^{-\psi}_{i,t+1} + b(1-p_t)\hat{a}^{-\gamma}_{i,t+1} \right) \Big] = 0 \tag{2.C.1}$$

$$(1-\delta)\hat{c}_{i,t}^{-\psi} = \delta\tilde{V}_{i,t}(\hat{s}_{i,t},\xi_{i,t},z_{i,t})^{\frac{\gamma-\psi}{1-\psi}} \times$$
$$\mathbb{E}_t\left[e^{-\gamma(\Delta w_{t+1}+\Delta f_{t+1})}\left(R_f+\xi_{i,t}R^e_{t+1}\right)\right. \tag{2.C.2}$$
$$\left.\times\left(p_t(1-\delta)\hat{V}_{i,t+1}^{\psi-\gamma}(\hat{a}_{i,t+1},z_{i,t+1})\hat{c}_{i,t+1}^{-\psi}+b(1-p_t)\hat{a}_{i,t+1}^{-\gamma}\right)\right]$$

Special cases.

Retirement. The problem's solution remains exactly the same except for the fact that $z_{i,t}, \alpha_i$ and β_i are not state variables anymore and that $\bar{L}_{i,t} = \bar{L}_{i,K}$, which is constant and with respect to, therefore, we do not need to compute expectations. Another difference is the retirement replacement ratio for labor income starting from $K+1$. Furthermore, because during retirement wages are not indexed anymore, it is not possible to scale the problem by the average wage as before. I overcome this issue by following Catherine (2021) and assuming that the average wage index remains constant after retirement.

Last period of life. Recall that in the last period of life T it holds $p_T = 0$ so that the objective function becomes:

$$\hat{V}_{i,T}(\hat{a}_{i,T},y_{i,K}) = \max_{\xi_{i,T},\hat{c}_{i,T},\hat{s}_{i,T}}\left\{(1-\delta)\hat{c}_{i,T}^{1-\psi}\right.$$
$$\left.+\delta\left[\mathbb{E}_T\left(b\hat{a}_{i,T+1}^{1-\gamma}e^{(\Delta w_{T+1}+\Delta f_{T+1})(1-\gamma)}\right)\right]^{\frac{1-\psi}{1-\gamma}}\right\}^{\frac{1}{1-\psi}}$$

subject to the same constraints as before. For the solution, it is again useful to define:

$$\tilde{V}_{i,T}(\hat{s}_{i,T},\xi_{i,T},y_{i,K}) = \left[\mathbb{E}_T\left(b\hat{a}_{i,T+1}^{1-\gamma}e^{(\Delta w_{T+1}+\Delta f_{T+1})(1-\gamma)}\right)\right]^{\frac{1-\psi}{1-\gamma}}$$

Proceeding as in the previous section, we get the following first order conditions:

$$\mathbb{E}_T\left[e^{-\gamma(\Delta w_{t+1}+\Delta f_{T+1})}R^e_{T+1}\hat{a}_{i,T+1}^{-\gamma}\right] = 0$$

$$(1-\delta)\hat{c}_{i,T}^{-\psi} = \delta \tilde{V}_{i,T}(\hat{s}_{i,T}, \xi_{i,T}, y_{i,K})^{\frac{\gamma-\psi}{1-\psi}} \\ \times \mathbb{E}_T\left[e^{-\gamma(\Delta w_{T+1}+\Delta f_{T+1})}\left(R_f + \xi_{i,T} R^e_{T+1}\right) b\hat{a}_{i,T+1}^{-\gamma}\right]$$

In the above equations we have assumed $b \neq 0$. If $b = 0$, the problem becomes very simple:

$$\hat{V}_{i,T}(\hat{a}_{i,T}, y_{i,K}) = \max_{\xi_{i,T},\hat{c}_{i,T},\hat{s}_{i,T}} (1-\delta)^{\frac{1}{1-\psi}} \hat{c}_{i,T}$$

with the same constraints as before. The trivial optimal policies are thus $\hat{c}_{i,T} = \hat{a}_{i,T}$ and $\hat{s}_{i,T} = 0$.

Solution algorithm. I will outline the solution algorithm for the most general case, special cases can be included as straightforward extensions. The model is solved by backward induction and the endogenous grid point method with the following procedure:

1. Use the terminal condition (2.2) to solve for the the value function and optimal policies at $T+1$;

2. For each $t \in [K+1, T]$ solve for the optimal policies and value function as follows:

    - For each point in the grid for $\bar{L}$ and for each for each point in the grid for $\hat{s}$ compute:

        (a) Optimal risky share $\xi_{i,t}$ in the case of participation and of non-participation. Recall that equation (2.C.1) solves $\frac{\partial \tilde{V}}{\partial \xi} = 0$. If $\frac{\partial \tilde{V}}{\partial \xi} > 0$ if $\xi = 1$ then set the optimal risky share to 1 while if $\frac{\partial \tilde{V}}{\partial \xi} < 0$ if $\xi = 0$ then set the optimal risky share to 0. Otherwise, set the optimal risky share to the value that solves (2.C.1). In the case of non-participation the optimal risky share is trivially zero;

        (b) Optimal consumption $\hat{c}_{i,t}$ by solving equation (2.C.2);

        (c) Cash on hand at the beginning of the period from the normalized constraint $\hat{c}_{i,t} + \hat{s}_{i,t} + F_{i,t}\kappa_f / e^{\bar{f}(t)} = \hat{a}_{i,t}$;

(d) The value function by inserting the optimal policies just computed in the expression of the value function $\hat{V}_{i,t}$;

(e) Using the minimum value of cash on hand implied by the model find if the borrowing constraint binds. In the affirmative case add a point corresponding to this case;

(f) Linearly interpolate the value function and the optimal policies on the grid for cash on hand at the beginning of the period. Note that this requires finding the switching point between participation and non-participation on the cash on hand grid point and using the solution values for non-participation below that point and for participation above that point. using the optimal quantities for the case

3. For each $t \in [T_{\text{start}}, K]$ solve for the optimal policies and value function as follows:

   - For each point in the grid for β, for each point in the grid for α, for each point in the grid for z, for each point in the grid for $\bar{L}$ and for each for each point in the grid for $\hat{s}$: repeat the same steps (a)-(e) in the list above.

Interpolation. The solution procedure outlined in section 2.C will very often require to evaluate the value function and the consumption policy at points off the grid. This also applies to model simulation when evaluating the solved policies at the points of the simulated paths. As explained in section 2.C, I do not discretize the persistent component of idiosyncratic income, which implies that I need to interpolate these functions not only at points off the cash on hand grid, the life-cycle parameters α, β grids, the average idiosyncratic income $\bar{L}$ grid, but also off the grid of persistent income. In other words, I need a multidimensional interpolation procedure over the $(\alpha, \beta, a, z, \bar{L})$ grid. This is achieved by multidimensional linear interpolation.

Computing expectations. In order to solve the model, it is necessary to compute expectations of some non trivial functions. In the most general case, I need to compute expectations with respect to the shocks r_1, r_2, ϕ, ε and ν.[38] To do that, I proceed as follows: (i) for all the possible combinations of grid values of these variables, I compute the value of the function (ii) I multiply it by the probability of that particular combination of values (iii) once I have done this for all the possible combinations I sum up all the function values obtained. Note that the grid values and probabilities of the other shocks coincide with Gaussian quadrature nodes and weights[39], which enables me to compute expectations very accurately. Finally, remember that the distributions of r_1 is conditional on the realization or not of a stock market crash and, similarly, those of ε and ν on the realization of a tail event or not. This is taken into account simply by scaling the probability of the discretized conditional distributions of these variables by the probability of these events.

2.D Estimation

This section describes how I estimate the exogenous stochastic processes. For both aggregate variables and individual income process I follow the procedure outlined in Catherine (2021). I will now describe the part of the procedure that is common for both and then dedicate two specific paragraph for the peculiarities regarding each of the two. The numerical parameters chosen for the estimation procedure are reported in Table 2.C.1.

Let θ be the vector of parameters that has to be estimated. θ is chosen to minimize the following objective function:

$$\min_{\theta} \hat{m}(\theta)' W \hat{m}(\theta) \tag{2.D.3}$$

[38] Cases in which I do not need to take expectations with respect to one or more of these variables can be handled by the same procedure outlined here with straightforward modifications.

[39] See section 2.C.

where $\hat{m}(\theta)$ is a vector of moments that depends on the parameters to be estimated and W is a weighting matrix. The procedure involves a global and a local stage. In the global stage I compute the value of the objective function for N_{glo} combination of points for the elements of the vector θ. The combinations correspond to the first N_{glo} of a Sobol sequence. At the end of the global stage, the best - in the sense of providing the lowest values of the objective function - N_{loc} points pass to the local stage. In the local stage, for each of the N_{loc} points, equation (2.D.3) is solved for the minimum using the Nelder-Mead algorithm with starting guess each of such points. The minimum is then the vector of parameters among the N_{loc} local points that returns the lowest value of the objective function.

Aggregate processes. To estimate the stochastic process governing the aggregate variables in the model I simulate the process for T_{cal} periods and then I compute the difference between the model generated moments and the moments in the data. I discard the first T_{dis} points from moments computation and, in order to smooth the surface of the objective function, I simulate the process for N_{eco} economies and average moments across them. Letting m indicate a generic moment, $\hat{m}(\theta)$ is defined as follows:

$$\hat{m}(\theta) = \frac{m_{\text{data}} - m_{\text{simulated}}(\theta)}{m_{\text{data}}} \tag{2.D.4}$$

The weighting matrix W is a unitary diagonal matrix. The actual moments I target are described in the main text.

Individual income process. I closely follow Guvenen et al. (2014) and Catherine (2021) to estimate the stochastic process governing individual income. Specifically, I simulate the income histories of 68 cohorts, the first starting in 1944 and the last in 2011 assuming that the persistent component is zero at the beginning and that the model economy is subject to the same aggregate wage shocks as in the data. Each cohort is made of N_{cal} individuals. Note that this implies having a constant age structure between 1979 and 2011.[40] Also in this case, to smooth the surface of the objective function I simulate N_{eco} economies and average moments across them. As described in the main text, the first 155 moments I match are standard deviation and Kelly's skewness of earnings growth at different time horizons. For these moments, the function $\hat{m}_t(\theta)$ is defined as follows:

$$\hat{m}_t(\theta) = \frac{m_{t,\text{data}} - m_{t,\text{simulated}}(\theta)}{\overline{m}_{\text{data}}} \tag{2.D.5}$$

where $\overline{m}_{\text{data}}$ is the time-series average of the absolute value of the moment under scrutiny. For the 36 within-cohort variances, instead, I use the same formula as in equation (2.D.4). The weighting matrix W is a unitary diagonal matrix that assigns equal weights to all moments.

[40]The moments provided by Guvenen et al. (2014) refer to individuals between 25 and 60.

Chapter 3

Preference heterogeneity and portfolio choices over the wealth distribution

This paper has been jointly written with Markus Kondziella and Zoltán Rácz. We are grateful to Timo Boppart, John Hassler, Per Krusell, Kieran Larkin, Lars Ljungqvist, Kurt Mitman and Paolo Sodini for their advice and support. We also thank Tobias Broer, Karin Kinnerud, Richard Foltyn and seminar participants at the IIES Macro Group for useful feedback and comments.

3.1 Introduction

Increased availability of data on portfolio holdings and wealth at the individual level has spurred growth in the literature on portfolio choice and capital income returns. On the empirical side, Bach et al. (2020) and Fagereng et al. (2020) study how portfolio choice and return heterogeneity vary over the wealth distribution. On the theoretical side, Benhabib et al. (2011) and Hubmer et al. (2021) emphasize the importance of return heterogeneity in explaining wealth accumulation over time at the individual level and wealth inequality in the cross-section.

Despite these recent advancements, explaining individuals' portfolio choices and cross-sectional wealth inequality remain two challenging issues in household finance and macroeconomics, respectively. In this paper, we show that connecting the two literatures by introducing a macroeconomic angle to the recent empirical findings in the household finance literature can help to address both issues simultaneously. Specifically, we extend an otherwise standard incomplete markets model along several dimensions to generate portfolio choice patterns consistent with the empirical findings and show that such extensions improve the match of the cross-sectional wealth distribution in the data, particularly at the very top.

The core of our framework is a Bewley model, the workhorse for studying the interplay between the wealth distribution and macroeconomic aggregates. We add to the standard setting endogenous portfolio choice, a non-normal return process, cyclical skewness in labor income shocks, Epstein-Zin preferences and preference heterogeneity. The latter includes heterogeneity in individuals' time preference rate (TPR), risk aversion and ability or inclination towards portfolio diversification.[1] The result is a hybrid between Bewley-type and portfolio choice models in the household finance literature.

[1]Our specification allows also for heterogeneity in the elasticity of intertemporal substitution (EIS). However, as discussed below in Section 3.3, in this paper we abstract from that because our framework does not allow to identify well this parameter.

We estimate the unobserved model parameters governing the heterogeneity in preferences to match the increasing risky share, participation rate and share of idiosyncratic return risk over the wealth distribution documented in Bach et al. (2020). As an outcome of the estimation, the economy is populated by the following two types: one type features a relatively high risk aversion and impatience (parameter values commonly used in the household finance literature), whereas the other type is characterized by lower risk aversion and impatience (parameter values commonly used in the macroeconomics literature). The former type also features a higher preference for portfolio diversification than the latter.

The combination of individuals of both types delivers portfolio choice patterns that closely match the patterns in the data. The aforementioned introduction of preference heterogeneity and the rich stochastic process governing income and returns - which we borrow from Catherine (2021)[2] - are crucial to generate the increasing relation between the risky share and wealth quantiles. In particular, the income process is relevant to explain the risky share towards the bottom of the wealth distribution and the preference heterogeneity towards the top. Intuitively, human capital is a higher share of net worth at the bottom, which makes the stochastic properties of income matter more than at the top, where, instead, preference parameters become the primary determinant of portfolio choices.

These two channels ensure a positive correlation between risky shares and wealth over the whole distribution. Towards the bottom, this trend follows from the hesitancy of asset-poor individuals to invest in the stock market because of the riskiness of their labor income. At the top, instead, such relationship is obtained only in combination

[2]While we take from his paper most of the elements governing the joint stochastic process of income and returns, it is important to state that we use them for different purposes than his paper. Indeed, rather than explaining portfolio choices over the life-cycle dimension, we are interested in understanding them over the wealth distribution and how they affect inequality through their impact on the formation of individual returns.

with preference heterogeneity. Intuitively, the optimal risky share of rich agents converges to the constant in Merton (1969), as high wealth essentially protects the individual from non-linear features in our model. If all agents share identical preference parameters, generating the empirically substantial positive relationship between wealth and risky share is therefore impossible. However, preference heterogeneity causes less risk-averse agents with a higher risky share to endogenously end up on the top of the wealth distribution generating a positive relationship between risky shares and wealth in the cross-section. In addition, because these individuals are characterized by lower portfolio diversification, we also match the higher share of idiosyncratic variance at the top. Together with the fact that their higher degree of patience ensures that less risk-averse and less diversified individuals endogenously end up at the top of the distribution, these mechanisms also increase wealth inequality through higher expected returns among the richest.

Finally, to gauge the relative importance of the different elements in our framework, we compare the results in our benchmark model with counterfactuals in which we shut down different components one at a time. More in detail, we solve a version with homogeneous preferences, one with heterogeneity in just the time preference rate and another in just risk aversion, a version without endogenous portfolio choice, one without idiosyncratic returns and one without skewness in labor income and return shocks. Except for the case without idiosyncratic returns, in which (in line with Hubmer et al., 2021) we find relatively small changes, in all the other cases either the match of the portfolio schedules or of wealth inequality or of both is worsened.

Related literature. This paper contributes to both the household finance literature and to the macroeconomics literature on wealth inequality.

Our main contribution to the former is capturing endogenously realistic portfolio choices over the wealth distribution as documented in

Swedish registry data by Bach et al. (2020). In doing that, our paper relates to the literature studying the interplay between income, preference heterogeneity and portfolio choices.

Both theoretical and empirical studies show that labor income is a determinant of portfolio choice. For example, the persistent component of labor income is linked to human capital, and from theory starting with Merton (1969) we know the latter influences participation and the risky share. Fagereng et al. (2017) and Chang et al. (2018) further emphasize that the riskiness of labor income influences portfolio choice (the riskier labor income, the lower the risky share). Using Swedish registry data Catherine et al. (2021) show that workers facing higher cyclical skewness display lower risky shares. In line with these findings - and those in Guvenen et al. (2014) - we follow Catherine (2021) and include an income process in our model that features skewness of idiosyncratic income shocks that is linked to movements in aggregate returns.

While including this central feature of Catherine (2021), we also extend his setting to allow a rich set of parameters governing preference heterogeneity.[3] Thus, our paper also relates to the literature studying the role of the latter in portfolio choice models (e.g., Vestman, 2018).[4] In addition, as we use our framework to structurally estimate the parameters governing preference heterogeneity, we also relate to the emerging household finance literature in this area (Calvet et al., 2021).

Within the literature on wealth inequality, several studies emphasize that capturing return heterogeneity is a crucial component to match the shape of the wealth distribution. Benhabib et al. (2011) show analytically that the introduction of stochastic idiosyncratic returns

[3] Also differently from his framework, we allow for idiosyncratic return shocks and use an infinite horizon setup.

[4] Vestman (2018) investigates the effects of joint heterogeneity in risk aversion, EIS and participation cost on stock market participation patterns and their connection with home ownership. However, he does not consider the risky share, which is the main focus of this paper. Furthermore, contrary to Vestman (2018), instead of presetting preference parameters, we are estimating them.

implies a wealth distribution that is Pareto. In their model, wealth accumulation and decumulation occur randomly. Benhabib et al. (2015), Nirei and Aoki (2016) and Gabaix et al. (2016) are further examples that introduce heterogeneous returns to the consumption-savings decision. Benhabib et al. (2019) quantitatively show that heterogeneous returns jointly with savings and bequests behavior are crucial elements for top wealth inequality and to explain social mobility. Hubmer et al. (2021) study plausible explanations for the increase in wealth inequality in the U.S. Heterogeneity in asset returns turns out to be key to account for the dynamics in wealth inequality. The return on assets is modelled as an increasing function of wealth plus an idiosyncratic shock. Thus, individuals end up with different returns both because they have different wealth levels (which can, potentially, be controlled through the savings decision) and because of randomness. That returns on assets are increasing in wealth can be interpreted as reduced-form portfolio choice that is consistent with the results by Bach et al. (2020) found in Swedish registry data.

Despite the fact that portfolio choice is a crucial component to generating individual returns, the above papers take shortcuts in obtaining return heterogeneity. In order to take the driver of return heterogeneity into account in models analysing the wealth distribution, it is of first-order importance to endogenize an individual's investment decision between different kinds of assets. Our contribution to this literature is, therefore, adding realistic endogenous portfolio choices to this class of models, and showing that the latter is crucial to capture wealth inequality. In doing that, we also try to connect this research area with the household finance literature described above.

Finally, as one important element in our framework to achieve a good match of the wealth distribution in the data is preference heterogeneity, we also relate to the papers emphasizing the role of the latter for inequality (see De Nardi and Fella, 2017, for a review). The new element in our paper is considering a richer structure compared to what has been done so far. In particular, our paper explores the consequences of introducing heterogeneity in impatience, risk-aversion and lack of diversification[5] - and of allowing correlations between them - on the wealth distribution through their impact on both agents' optimal consumption-savings and portfolio choices. As we will see below, we find this to be an important element to fit the data and to investigate the relative importance of different channels.

The paper is structured as follows. Section 3.2 outlines the model, section 3.3 describes our calibration procedure, section 3.4 presents results on our benchmark specification, section 3.5 investigates counterfactuals and section 3.6 concludes.

3.2 Model

Agents and preferences. The economy is populated by a continuum of infinitely lived ex-ante identical individuals deriving utility from consumption ($c_{i,t}$) through Epstein-Zin preferences. Agents differ in terms of preference parameters. We capture this heterogeneity with an individual-specific preference state ($\theta_{i,t}$) which, in turn, determines impatience (δ), risk aversion (γ), the inverse of the elasticity of intertemporal substitution (ψ) and the lack of diversification[6] (ζ). Preferences are, then, given by the following expression:

$$U_{i,t} = \left[(1-\delta(\theta_{i,t}))c_{i,t}^{1-\psi(\theta_{i,t})} + \delta(\theta_{i,t}) \left(\mathbb{E}_t U_{i,t+1}^{1-\gamma(\theta_{i,t})} \right)^{\frac{1-\psi(\theta_{i,t})}{1-\gamma(\theta_{i,t})}} \right]^{\frac{1}{1-\psi(\theta_{i,t})}}$$

Financial assets. Agents can invest in two financial assets, one risky with time-varying individual-specific gross return $R_{i,t+1}$ and one safe

[5]Including lack of diversification as a preference parameter allows us to match in which part of the wealth distribution idiosyncratic return variance is more important, and hence we can investigate the relative importance of idiosyncratic return shocks realistically.

[6]The next paragraph provides a detailed explanation of lack of diversification.

with constant gross return R^f. Letting small letters indicate log returns, $r_{i,t+1}$ is equal to:

$$r_{i,t+1} = r_{1,t+1} + r_{2,t+1} + \eta_{i,t+1} - m \tag{3.1}$$

The effective return individual i gets by investing in the risky asset is the sum of two systematic components, one co-varying with labor market conditions (r_1) and one that does not (r_2), of an idiosyncratic component (η) and is net of management cost m, that is thus paid conditional on holding the risky asset. The systematic components are modeled as in Catherine (2021). Specifically, to take into account stock market crashes, r_1 is distributed as a mixture of Normals:

$$r_{1,t+1} = \begin{cases} \underline{r}_{1,t+1} \overset{i.i.d.}{\sim} \mathcal{N}\left(\underline{\mu}_r, \sigma^2_{r_1}\right) & \text{w.p.} \quad p_r \\ \overline{r}_{1,t+1} \overset{i.i.d.}{\sim} \mathcal{N}\left(\overline{\mu}_r, \sigma^2_{r_1}\right) & \text{w.p.} \quad 1 - p_r \end{cases} \tag{3.2}$$

Without loss of generality, we interpret p_r as the probability of stock market crashes and $\underline{\mu}_r$ the expected log return during these periods. Similarly, $1 - p_r$ is the probability of normal periods and $\overline{\mu}_r$ the corresponding average log return. The other systematic component, r_2, is drawn from a Normal distribution:

$$r_{2,t+1} \overset{i.i.d.}{\sim} \mathcal{N}\left(0, \sigma^2_{r_2}\right)$$

Finally, the idiosyncratic component, $\eta_{i,t+1}$, is modeled as follows:

$$\eta_{i,t+1} \overset{i.i.d.}{\sim} \mathcal{N}\left(-\frac{\sigma^2_{ir}}{2}, \sigma^2_{ir}\right)$$

where $\sigma_{ir} = \sigma_r \zeta(\theta_{i,t})$, with σ_r being the standard deviation of the systematic part of the log return and the preference parameter $\zeta(\theta_{i,t})$ governing the share of idiosyncratic risk in total portfolio volatility. Note that this specification ensures that idiosyncratic risk is not priced since the idiosyncratic part does not affect the mean return.[7]

Introducing the idiosyncratic component enables us to understand the relative importance of systematic and idiosyncratic return shocks. However, as standard portfolio choice models imply that everyone should invest in some efficient mixture of riskless and risky assets, it is far from straightforward to represent lack of diversification in a framework otherwise based on optimizing behavior.

Our modeling choice implies the following: rather than having access to the same risky asset, each individual rationally invests in her own risky asset, which has identical expected excess return as the market, but additional preference-dependent idiosyncratic risk. While guaranteeing that idiosyncratic risk is not priced, this strategy also implies - in line with the empirical findings in Calvet et al. (2007) - that agents worse at diversifying will, everything else equal, optimally choose a lower risky share, and vice versa.

However, linking the share of idiosyncratic risk to a stable preference type also effectively restricts the domain of portfolio composition decisions. In other words, we do not model how lack of diversification arises (e.g., financial knowledge, overconfidence, reliance on private equity) but, rather, capture in reduced form that agents' ability or desire to diversify is limited and that they optimally decide how to allocate their wealth given this constraint. Thus, our approach lies between a completely micro-founded, realistic model of portfolio choice in the presence of a menu of different risky assets, and a framework in which the stochastic properties of returns over the wealth distribution are hard-wired (Hubmer et al., 2021).

Investing in the risky asset is subject to a fixed participation cost f that is paid in every period the agent chooses to hold that asset. Finally, individuals face a borrowing limit on their total savings proportional to the exogenously set parameter $\overline{s}$. The repayment rate per unit of borrowing is equal to the risk-free rate.

[7] For the exact formulas of the statistical moments of the full return process see Appendix 3.C.

Labor income process. We follow Catherine (2021) for modeling the stochastic process governing labor income. Let $y_{i,t}$ denote the residual of log individual earnings. We assume that $y_{i,t}$ is the sum of an aggregate component (w_t) and of two idiosyncratic components, one persistent ($z_{i,t}$) and one transitory ($\nu_{i,t}$):

$$y_{i,t} = w_t + z_{i,t} + \nu_{i,t} \tag{3.3}$$

The aggregate component follows a random walk with drift, driven by shocks to the market return through a parameter λ_{rw}:

$$w_t = g + w_{t-1} + \lambda_{rw} r_{1,t} + \phi_t \tag{3.4}$$

where $\phi_t \overset{i.i.d.}{\sim} \mathcal{N}\left(0, \sigma_\phi^2\right)$.

The persistent component is an AR(1) process:

$$z_{i,t} = \rho z_{i,t-1} + \varepsilon_{i,t} \tag{3.5}$$

with innovations drawn from a mixture of Normals:

$$\varepsilon_{i,t} = \begin{cases} \underline{\varepsilon}_{i,t} \overset{i.i.d.}{\sim} \mathcal{N}\left(\underline{\mu}_{\varepsilon,t}, \underline{\sigma}_\varepsilon^2\right) & \text{w.p.} \quad p_\varepsilon \\ \overline{\varepsilon}_{i,t} \overset{i.i.d.}{\sim} \mathcal{N}\left(\overline{\mu}_{\varepsilon,t}, \overline{\sigma}_\varepsilon^2\right) & \text{w.p.} \quad 1 - p_\varepsilon \end{cases} \tag{3.6}$$

Without loss of generality, we interpret p_ε as the probability of tail events and $\underline{\mu}_{\varepsilon,t}$, $\underline{\sigma}_{\varepsilon,t}$ the expected value and standard deviation of persistent income shocks during tail events, respectively. A similar interpretation holds for the parameters governing the distribution of normal events. To match the cyclicality of skewness, $\underline{\mu}_{\varepsilon,t}$ is defined as:

$$\underline{\mu}_{\varepsilon,t} = \mu_\varepsilon + \lambda_{\varepsilon w}(w_t - w_{t-1}) \tag{3.7}$$

Thus, tail events imply on average higher persistent shocks during expansions and vice versa during recessions. In addition, since these shocks have zero mean, it must hold:

$$p_\varepsilon \underline{\mu}_{\varepsilon,t} + (1 - p_\varepsilon)\overline{\mu}_{\varepsilon,t} = 0 \tag{3.8}$$

Finally, the transitory shock is Normally distributed, with variance depending on whether the persistent shock was drawn from the tail distribution or not:

$$\nu_{i,t} = \begin{cases} \underline{\nu}_{i,t} \overset{i.i.d.}{\sim} \mathcal{N}\left(0, \underline{\sigma}_\nu^2\right) & \text{if} \quad \varepsilon_{i,t} = \underline{\varepsilon}_{i,t} \\ \overline{\nu}_{i,t} \overset{i.i.d.}{\sim} \mathcal{N}\left(0, \overline{\sigma}_\nu^2\right) & \text{if} \quad \varepsilon_{i,t} = \overline{\varepsilon}_{i,t} \end{cases} \tag{3.9}$$

As discussed in detail by Catherine (2021), including countercyclical income risk enables us to obtain realistic portfolio choices over the life cycle. Intuitively, if adverse income shocks occur with greater probability when the stock market is low, agents with relatively high human capital (i.e., the young and the poor) will be more cautious when investing in risky financial instruments.

The reason for retaining this feature in our framework is two-fold. First, examining whether a mechanism allowing to match portfolio choices over age can also produce realistic risk-taking patterns over the wealth distribution is an interesting question. Second, by including a potential alternative channel of generating increasing risky shares over the wealth distribution relative to preference heterogeneity, we let the estimation decide whether the latter is crucial to match the desired patterns.

Safety net. To take into account how welfare programs potentially affect consumption, savings and portfolio choice, we follow Catherine (2021) who models the Supplemental Nutrition Assistance Program (sometimes called food stamp program). Specifically, working-age individuals with wealth holdings below 5% of the average national wage and earnings below 20% of the wage index receive 6% of the wage index minus 30% of their earnings as benefits. Mathematically:

$$b_{i,t} = \max\left\{0.06 \cdot \exp(w_t) - 0.3 \cdot \exp(y_{i,t}), 0\right\} \qquad \begin{array}{ll} \text{if} & a_{i,t} < 0.05 \cdot e^{w_t} \\ \text{and} & e^{y_{i,t}} < 0.2 \cdot e^{w_t} \end{array} \tag{3.10}$$

where $b_{i,t}$ denotes the benefits and $a_{i,t}$ cash-on-hand, which is defined in the next paragraph.

The optimization problem. At the beginning of each period t, the agent enters with given cash-on-hand $a_{i,t}$, persistent income $z_{i,t}$, preference state $\theta_{i,t}$, and aggregate income w_t. She then chooses how much to consume in the current period $c_{i,t}$, how much to save for the next period $s_{i,t}$, whether to hold risky assets $F_{i,t}$ (dummy equal to one if she participates) and, conditional on participation, the share of savings invested in risky assets $\xi_{i,t}$.

Let $\Xi_{i,t} := (a_{i,t}, z_{i,t}, \theta_{i,t}, w_t)$ denote the state, $R^f := \exp(r^f)$ the gross risk free return and $R^e_{i,t+1} := \exp(r_{i,t+1}) - R^f$ the excess return. Then the maximization problem of agent i is:

$$\begin{aligned} V(\Xi_{i,t}) = \max_{\{c_{i,t}, s_{i,t}, \xi_{i,t}, F_{i,t}\}} & \Big\{ (1-\delta(\theta_{i,t})) c_{i,t}^{1-\psi(\theta_{i,t})} \\ & + \delta(\theta_{i,t}) \left(\mathbb{E}_t \left[V(\Xi_{i,t+1})^{1-\gamma(\theta_{i,t})} \right] \right)^{\frac{1-\psi(\theta_{i,t})}{1-\gamma(\theta_{i,t})}} \Big\}^{\frac{1}{1-\psi(\theta_{i,t})}} \end{aligned} \tag{3.11}$$

subject to

$$c_{i,t} + s_{i,t} + F_{i,t} f = a_{i,t} \tag{3.12}$$

$$a_{i,t+1} = \left[R^f + \xi_{i,t} R^e_{i,t+1} \right] s_{i,t} + \exp(y_{i,t+1}) + b_{i,t+1} \tag{3.13}$$

$$s_{i,t} \geq \bar{s} \cdot \exp(w_t) \tag{3.14}$$

The borrowing constraint (3.14) varies over time through the dependence on the aggregate part of labor income w_t, which can be interpreted as this constraint becoming tighter in recessions and looser in expansions.[8] Appendix 3.B describes in detail how the model is solved numerically.

3.3 Estimation and calibration

The goal of the calibration is to deliver a parametrized model in line with novel empirical evidence on portfolio choice over the wealth distribution documented in Bach et al. (2020). To do so, we follow a two-step approach. First, we exogenously set the parameters governing the income and return processes. Second, we estimate the fixed participation cost and individuals' preference parameters. We describe in the following the details of the procedure adopted.

3.3.1 Exogenously set parameters

Since we model the stochastic processes governing income and returns as in Catherine (2021), we use the same parameter estimates reported in his paper. While an extensive explanation of the estimation procedure can be found there, we still provide a brief description of the approach. The parameters governing the aggregate processes $(\underline{\mu}_r, \overline{\mu}_r, \sigma_{r_1}, \sigma_{r_2}, p_r, \sigma_\phi, g, \lambda_{rw})$, are estimated by Simulated Method of Moments (SMM) to capture the joint dynamics of log yearly SP500 returns and aggregate wage log growth from US Social Security panel data on earnings by targeting mean, standard deviation, third and fourth standardized moments (skewness and kurtosis) and the correlation between these two series. Estimation of the stochastic process for individual income requires, instead, to find values for $(p_\varepsilon, \mu_\varepsilon, \lambda_{\varepsilon w}, \underline{\sigma}_\varepsilon, \overline{\sigma}_\varepsilon, \underline{\sigma}_\nu, \overline{\sigma}_\nu, \rho)$. To do so, SMM is used again targeting the time series between 1978 and 2010 of the standard deviation of log earnings growth at the one- and five-year horizons, Kelly's skewness of log earnings growth at the one-, three- and five-year horizons (from

[8]This assumption also makes the value function homogeneous with respect to w_t, which allows us to reduce the dimensionality of the problem by one.

Guvenen et al., 2014) and the within-cohort variance of log earnings for ages between 25 and 60 (from Guvenen et al., 2021).

The risk-free rate r^f is set as in Catherine (2021) to 0.02, which is a standard value in the literature (e.g., Cocco et al., 2005; Gomes and Michaelides, 2005). Finally, no borrowing is allowed so $\bar{s}$ is set to zero. Table 3.1 summarizes all these parameter choices.

3.3.2 Estimated parameters

The main contribution of our estimation exercise is to obtain structural estimates of individuals' preference parameters - and of their heterogeneity - from their observed portfolio choices over the wealth distribution. To achieve this, we assume that the economy is populated by two types (i.e., the support of $\theta_{i,t}$ has two states).[9]

It is worth noticing that, as θ is a vector including time preference rate (δ), risk aversion (γ), inverse EIS (ψ), and lack of diversification (ζ), our model enables the investigation of potential heterogeneities across all these dimensions simultaneously. Nevertheless, as it cannot be well identified in our framework, in all the results reported from here onwards, ψ is set to unity for all types. Indeed, while the different role of ψ from that of risk aversion γ is discerned through the adoption of Epstein-Zin preferences, as highlighted by Aguiar et al. (2021), joint identification of EIS and time preference rate is problematic in a model without liquidity differences across assets.[10] In any case, we show in the results section below that heterogeneity in ψ is not necessary to match our targets.

[9]In our current specification, agents' preference types are assumed to remain fixed over time, as this reduces considerably the computational time. We have experimented with an extension of the model allowing transitions between types according to a Markov process in which the transition probabilities are jointly estimated with the other parameters and did not find not significant differences.

[10]The main idea behind identification in their framework is that low EIS agents care more about consumption smoothing and so invest more in liquid assets. We plan to include the joint estimation of the EIS with the remaining preference parameters in future versions of this paper.

Summing up, eight parameters are estimated in total: six preference parameters (δ, γ, ζ, for each type), the share of individuals of the first type, and the fixed participation cost.

Targets. The main targets of our estimation are portfolio choice patterns over the wealth distribution. To compute them, we rely on the data from Bach et al. (2020), compiled from Swedish administrative sources covering earnings and wealth holdings of all Swedish residents and on the figures already available in their appendix. Our data spans the period 2000–2007. While a detailed description can be found in their paper, for our purposes it is worth reminding that wealth holdings cover cash, pension wealth, financial securities (including funds, stocks, derivatives, and bonds), private equity, real estate wealth, and debt. These data are then aggregated at the household level using household identifiers. The measure of wealth we will refer to throughout the paper is net wealth, defined - as they do - as the sum of all wealth holdings within the household minus debt.

When deciding how to allocate their savings, in our model individuals choose between a safe and a single "composite" risky asset. To map excess returns, participation and the share of idiosyncratic risk by asset type in Bach et al. (2020), into those of a composite risky asset, we proceed as follows. First, we classify the different assets into safe and risky: cash, money market funds, pension wealth and residential real estate belong to the former group, while all other securities, private equity and commercial real estate to the latter. We then define the participation rate for the composite asset as the share of people in each wealth quantile holding any risky asset classified as such according to the method just described.[11] To obtain the excess return and the portfolio share of idiosyncratic risk in each wealth quantile, we multiply,

[11]We kindly thank Paolo Sodini for sharing the moments on participation consistent with the classification of assets into safe and risky applied in our paper, as well as the moments on the Swedish wealth distribution that we compare our estimated model to in section 3.4.2.

instead, the reported share of wealth invested in each asset type by, respectively, the expected excess returns and the share of idiosyncratic risk for that particular asset. We further rescale these excess returns by the average yearly excess return of the SIXRX Swedish equity index, that is is 8.7% over the 1983-2016 period (Bach et al., 2020). This transforms excess returns into the implicit unconditional risky share invested in the risky asset and eases comparison with other studies in the literature.

Figure 3.1 displays the resulting schedules of the unconditional risky share, participation and share of idiosyncratic risk over the wealth distribution. As in Bach et al. (2020), wealthier households are more likely to hold risky assets, invest a higher share of their wealth in those risky assets and load their portfolios with more idiosyncratic risk than poorer households. These three schedules constitute our calibration targets, together with the ratio of aggregate wealth to income - which in Sweden is equal to four as reported by Bach et al. (2018) - for a total of 49 moments.

Estimation results. The SMM estimation procedure comprises a global and a local stage.[12] In the global stage, we generate 1,000 parameter vectors from a Sobol sequence.[13] For every parameter vector Φ, we solve and simulate the model and then evaluate:

$$d(\Phi)'\Omega d(\Phi), \tag{3.15}$$

where $d(\Phi)$ is a vector containing the implied deviations of the model moments from their targets in the data and Ω is a diagonal weight-

[12] To preserve computational feasibility, for the estimation procedure we adopt smaller grids than those reported in Appendix 3.B, which we use in the results section. Specifically, we use 3 quadrature nodes, 75 points for the cash-on-hand grid, and 15 points for the grid of income's persistent component.

[13] Again for computational feasibility, guided by the arguments that will be outlined below when describing the results, we restrict the global stage of the estimation to search in regions of the parameter space in which the predominant type has jointly lower time preference rate, higher risk aversion, and higher portfolio diversification. The local stage, instead, is not bounded by this constraint.

ing matrix. The deviations from the targets in $d(\Phi)$ are computed as a percentage deviation for the wealth-to-income ratio and relative to the average over the wealth distribution for the remaining targets. The weighting matrix puts 50% of the weight on the wealth-to-income ratio and splits the remaining 50% equally between the schedules of the unconditional risky share, participation, and share of idiosyncratic variance over the wealth distribution. We choose the parameter vector from the Sobol sequence that minimizes equation (3.15) and proceed to the local step. At this stage, we take the candidate from the first step as an initial starting point and perform a local optimization using the Nelder-Mead algorithm to minimize again equation (3.15).

The estimated parameters are presented in Table 3.1. Type-two individuals discount the future less strongly than type-one agents (δ of 0.96 vs. 0.87), are less risk averse (γ of 1.37 vs. 10.31), and seek a lower degree of portfolio diversification. In particular, the estimated values for portfolio diversification imply that the share of idiosyncratic variance in return variance for type-two individuals is 57% ($\zeta = 1.08$), whereas it is 28% for type-one agents ($\zeta = 0.59$).[14] As illustrated by the figures just reported, the results imply a very stark separation across the two types in terms of preferences. In particular, the lower risk aversion and diversification found for type-two agents resemble common anecdotal traits among entrepreneurs.[15]

[14]The share of idiosyncratic variance in total return variance can be computed using the formula for return variance reported in Appendix 3.C.

[15]There is a vast literature on entrepreneurship and wealth inequality (see De Nardi and Fella, 2017, for a review) emphasizing the tension between individual ability and borrowing frictions and its impact on agents' savings behavior, an element which is not present in our model but that might have been captured by our estimation procedure through preference parameters.

Preference parameters				
		Type 1	*Type 2*	
δ	time preference rate	0.87	0.96	estimated
γ	risk aversion	10.31	1.37	estimated
ψ	inverse EIS	1.0	1.0	preset
ζ	lack of diversification	0.59	1.08	estimated
	share of individuals	0.96	0.04	estimated
Participation and management costs, borrowing limit				
f	fixed participation cost		0.001	estimated
m	management fee		0.01	Catherine (2021)
$\overline{s}$	borrowing limit		0	preset
Returns				
r^f	risk-free rate		0.02	
$\underline{\mu}_r$	mean syst. return crashes		-0.245	
$\overline{\mu}_r$	mean syst. return normal times		0.115	Catherine (2021)
σ_{r_1}	cond. st. dev. syst. return, part linked to w		0.077	
p_r	probability crashes		0.146	
σ_{r_2}	st. dev. syst. return, part not linked to w		0.114	
Income				
g	drift aggregate wage growth		0.008	
λ_{rw}	sensitivity aggregate wage growth to return		0.161	
σ_ϕ	st.dev. aggregate wage growth shock		0.017	
ρ	autocorrelation persistent component		0.967	
μ_ε	constant mean persistent shock, tail		-0.086	
$\lambda_{\varepsilon w}$	sensitivity mean perm. shock to Δw, tail		4.291	Catherine (2021)
p_ε	probability tail events		0.136	
$\underline{\sigma}_\varepsilon$	st.dev. persistent shock, tail		0.562	
$\overline{\sigma}_\varepsilon$	st.dev. persistent shock, non-tail		0.037	
$\underline{\sigma}_\nu$	st.dev. transitory shock, tail		0.895	
$\overline{\sigma}_\nu$	st.dev. transitory shock, non-tail		0.089	

Table 3.1: Model parameters values.

Type-one individuals are predominant in this economy, as they make up 96% of the total population. This is an interesting result for two reasons. First, it is surprising that, while the preference parameters of type-two agents are more in line with those usually adopted in the macro literature on wealth inequality, the majority of individuals are of type-one, which is characterized by lower time preference rate and high risk aversion, a combination more often used in the household finance literature. Second, despite what just noted, it indicates that only a small fraction of type-two agents is needed to replicate portfolio choice patterns over wealth in the data (and, as we will see below, also wealth inequality).

Finally, the stock market participation cost f is estimated at 0.001, which is roughly 0.05% of the average yearly income and is smaller than the values typically found in the literature (e.g., Vissing-Jorgensen, 2003).

3.4 Model fit

3.4.1 Targeted moments

We begin by assessing how well the model, which we refer to as our *benchmark model* in the following sections, matches the targeted moments. The wealth-to-income ratio is 4.09, which is close to 4, the value in the data. The unconditional risky share (defined as the individual portfolio's expected excess return over the market's), the share of idiosyncratic risk in the total variance of the individual portfolio, and the participation rate over the wealth distribution are reported in Figure 3.1 against their empirical counterparts.

Overall, the calibration procedure matches well the portfolio choices over the wealth distribution, even though it slightly undershoots the participation rate in the bottom quantiles.

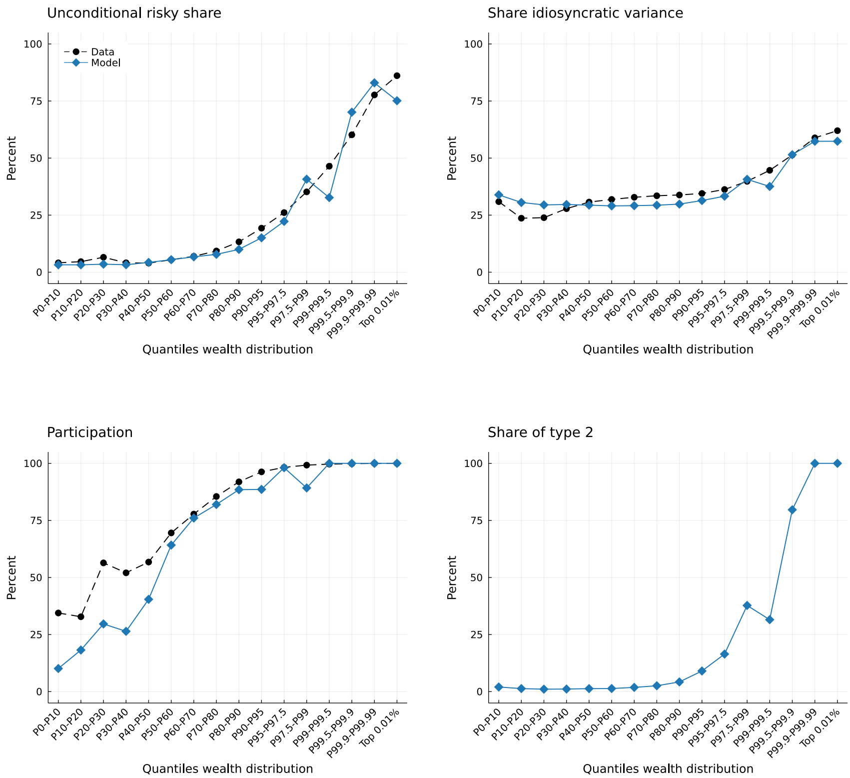


Figure 3.1: Policies (model vs. data) and share of Type 2 individuals over the wealth distribution.

Why does preference heterogeneity enable us to closely replicate the empirical patterns? The outcome of the calibration implies that the economy is predominantly populated by type-one individuals, distinguished by higher impatience, higher risk aversion, and lower share of idiosyncratic risk in their portfolios. Only 4% are type-two agents, who feature opposite characteristics. However, as shown in the bottom right panel of Figure 3.1, mainly thanks to their higher δ parameter, in equilibrium individuals of the latter type endogenously concentrate at the top of the wealth distribution.[16] As in the model portfolio choice patterns over the wealth distribution are largely determined by the relative

[16]Figure 3.A.1 reports the mass of the two types in levels.

share of types, the increasing number of type-two individuals - whose estimated preference parameters imply a high risky share and high idiosyncratic risk at the same time - over wealth allows the framework to reproduce the trends in the data.

It is also worth noticing that the estimated participation fixed cost, 0.05% of average yearly income, is very small compared to other (often unrealistically large) values used in the literature. The fact that more risk-averse individuals mainly populate the bottom of the wealth distribution - where participation is not an obviously optimal choice - is the reason why our model can match participation and portfolio choices with a low fixed cost. Indeed, high-risk aversion combined with the effect of countercyclical income risk, implies that even a small value of this parameter has a sufficient deterring effect on stock market entry.[17]

3.4.2 Untargeted moments: wealth distribution

Table 3.2 presents the model's performance in matching the share of total wealth held by different quantiles of the wealth distribution. We compare our results with empirical values computed by Krueger et al. (2016) using PSID (2006) and SCF (2007) data and, since we use portfolio choice moments from Swedish administrative data to estimate the model, with corresponding measures of the Swedish wealth distribution. Furthermore, to check how our framework compares to a state-of-the-art model of wealth inequality without portfolio choice, we add a column with the values generated by the benchmark model in Krueger et al. (2016).

[17]This mechanism is missing in a representative agent model in which, by construction, a high fixed cost is required to discourage the average agent from participating.

Share held by (%):	Benchmark	Krueger et al. (2016)	PSID (2006)	SCF (2007)	Sweden (2000-2007)
Q1	2.2	0.3	-0.9	-0.2	-1.1
Q2	5.6	1.2	0.8	1.2	2.8
Q3	9.2	4.7	4.4	4.6	8.7
Q4	15.5	16.0	13.0	11.9	19.4
Q5	67.5	77.8	82.7	82.5	70.2
90-95 %	11.5	17.9	13.7	11.1	13.4
95-99 %	22.5	26.0	22.8	25.3	17.9
Top 1 %	20.1	14.2	30.9	33.5	21.3
Wealth Gini	0.64	0.77	0.77	0.78	0.69

Table 3.2: Share of wealth held by people in different quantiles of the wealth distribution: benchmark model vs. data and a state-of-the-art model of wealth inequality.

The model matches well the share of wealth held by the 90-95%, 95-99%, and top 1% groups in the US. In the first two cases, it gets closer than Krueger et al. (2016) who, instead, overshoot the actual values. In the last case, despite being still very far from the corresponding figure in the data, compared to theirs, our model is able to generate a six percentage points higher share of wealth held. Remarkably, our framework delivers an even better match of top wealth inequality in Sweden[18], except for the share held by the 95-99% group, which is slightly higher in the model.

When looking at the distribution as a whole, instead, the performance is less satisfactory. In particular, the first three quintiles hold too much wealth compared to the data, and, as a consequence, the last quintile holds too little. This translates into lower Gini coefficients for wealth inequality than the actual values. Allowing for borrowing - and thus for agents to have negative wealth - might attenuate this issue, which is particularly relevant for the first quintile.[19]

[18]This is likely related to the fact that we estimate the preference parameters using portfolio choices over wealth from Swedish data.

[19]We plan to extend the model towards this direction in future research.

3.5 Counterfactuals

After having presented the fit of our benchmark specification, in this section we investigate the role of different model components in matching the targeted moments and generating a realistic wealth distribution. To this end, we shut down different features of our benchmark model and quantify the counterfactual predictions.

3.5.1 Homogeneous preferences

One of the main novelties of this paper is introducing rich heterogeneity in agents' preferences. In the following, we argue that this feature of our framework is crucial for explaining the targeted moments on portfolio choice over the wealth distribution shown in Figure 3.1. To this end, we re-estimate the model restricting preference parameters to be identical for both types.[20] Table 3.3 reports the parameter estimates for this case. With only one type, the values are in between the figures obtained in the benchmark case, as this minimizes the differences at the extrema of the schedules. This is also clearly visible from Figure 3.2, which shows the unconditional risky share (left panel) and the share of idiosyncratic variance (right panel) over the wealth distribution for this specification and the benchmark model. Notably, the estimated fixed cost is higher than in the benchmark case: as there is just one value for risk aversion, agents at the bottom need to be discouraged more from participating.[21]

The first argument for why the model without preference heterogeneity cannot deliver the empirical patterns concerns the schedule of the unconditional risky share over the wealth distribution. In the data - see Figure 3.1 - the risky share is increasing throughout the distribution, a feature which is captured by our model with preference heterogeneity. With only one type, instead, the unconditional risky share is

[20] To be precise, we re-estimate δ, γ, ζ and the participation cost f while targeting the same moments as for the benchmark model.

[21] Figure 3.A.2 reports the schedule of participation over wealth.

flat at zero for the first four deciles and increases up to roughly 25% between the fourth and seventh decile. In contrast to the empirical patterns, risky shares are constant or even decline in wealth for the top three deciles of the distribution.

Apart from the participation fixed cost, that risky shares initially increase in wealth is due to the forces highlighted in Catherine (2021). Despite their high human capital-to-wealth ratio, asset-poor individuals choose not to invest in the stock market due to the riskiness of their labor income. The strength of this effect vanishes as individuals become richer, but, at the same time, their human capital-to-wealth ratio declines. The risky share, therefore, plateaus before converging back to Merton's constant (Merton, 1969).

Despite the mechanism just described, the benchmark model with preference heterogeneity delivers increasing risky shares even at the top of the wealth distribution thanks to compositional effects. Indeed, due to their lower risk aversion, the risky share of type-two individuals converges to a higher constant than for type-one agents. As the share of type two individuals rises at the top, the risky share increases.

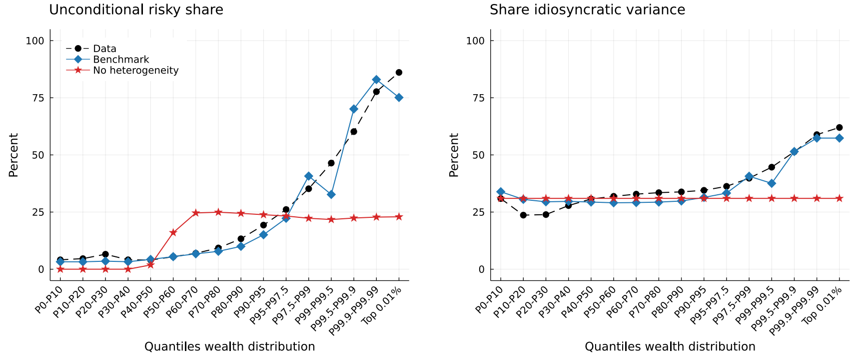


Figure 3.2: Policies over the wealth distribution in the model without preference heterogeneity, compared to the benchmark and data.

The second argument is related to the schedule of the share of idiosyncratic risk in total return risk over the wealth distribution. Figure

3.1 shows that (as in Bach et al. (2020)) the share of idiosyncratic risk is increasing in wealth, i.e., wealthier households hold relatively more idiosyncratic risk in their risky portfolios.

In our setting, the variance of the idiosyncratic return component is governed by ζ, i.e., individuals with the same ζ face the same idiosyncratic risk. Heterogeneity in the share of idiosyncratic return variance, therefore, arises only through differences in ζ across agents.

Compositional effects are again crucial for our benchmark model to replicate the empirical patterns for this schedule. Type-one individuals with relatively low ζ (and thus low idiosyncratic risk) mostly populate the bottom of the wealth distribution, whereas type-two agents with relatively high ζ (and thus high idiosyncratic risk) endogenously end up at the top. As a result, the share of idiosyncratic risk is increasing over the wealth distribution. As illustrated in the right panel of Figure 3.2, without ζ heterogeneity, the same quantity is constant.

The previous two points highlight the role of heterogeneity in risk aversion γ and portfolio diversification ζ. In both cases, we described that compositional effects due to the endogenous sorting of the two types over the wealth distribution were key to generating the increasing schedules of the risky share and of the share of idiosyncratic risk. To reinforce the argument, therefore, it is also important to highlight that attributing a higher degree of patience δ to type-two individuals (less risk averse and less diversified) and a lower one to type-one agents (risk averse and diversified) ensures that the former endogenously end up at the top of the wealth distribution.

Estimated value:		Benchmark	No het.	Only δ	Only γ	No idio. ret.	No skew.
Time preference rate, δ	*Type 1*	0.87	0.91	0.75	0.92	0.88	0.88
	Type 2	0.96		0.96	0.92	0.96	0.96
Risk aversion, γ	*Type 1*	10.31	6.54	5.42	4.34	9.22	18.42
	Type 2	1.37		5.42	6.31	1.20	1.28
Diversification, ζ	*Type 1*	0.59	0.63	0.67	0.76	0	0.64
	Type 2	1.08		0.67	0.76	0	1.09
Fixed cost, f		0.001	0.020	0.007	0.079	0.001	0.005
Share of Type 1		0.96	1	0.74	0.69	0.96	0.97

Table 3.3: Parameter estimates: benchmark model vs. alternative specifications. *"No het."* indicates the model without preference heterogeneity, *"Only δ"* the model with only heterogeneity in δ, *"Only γ"* the model with only heterogeneity in γ, *"No idio. ret."* the model without idiosyncratic returns and *"No skew."* the model without tail income shocks, stock market crashes and correlation between the income and return processes.

In addition to the effects on portfolio choice, shutting down preference heterogeneity has further implications for wealth inequality. It is known at least since Krusell and Smith (1998) that heterogeneity in patience (δ in our model) across individuals, can generate higher wealth inequality than the restricted case. The reason is that more patient individuals with higher saving rates are concentrated at the top of the wealth distribution in equilibrium, generating a longer right tail. The quantitative impact of preference heterogeneity on the wealth distribution is shown in Table 3.4. The Gini coefficient declines from 0.64 in the benchmark model to 0.61 in the model without preference heterogeneity, mainly due to a lower share of wealth held by the top quantile.

Share held by (%):	Benchmark	No het.	Only δ	Only γ	No port.	No idio. ret.	No skew.
Q1	2.2	1.9	0.5	1.2	2.4	2.2	2.3
Q2	5.6	5.4	1.5	4.0	6.5	5.7	5.9
Q3	9.2	10.0	2.9	7.7	10.8	9.6	9.7
Q4	15.5	17.9	10.2	16.1	18.0	16.2	16.2
Q5	67.5	64.8	85.0	71.0	62.3	66.3	65.8
90-95 %	11.5	12.8	17.6	14.0	12.3	12.0	11.2
95-99 %	22.5	20.2	28.4	22.9	18.8	21.6	22.2
Top 1 %	20.1	16.0	21.1	17.9	16.1	18.3	19.0
Wealth Gini	0.64	0.61	0.79	0.68	0.58	0.62	0.62

Table 3.4: Share of wealth held by people in different quantiles of the wealth distribution: benchmark model vs. alternative specifications. *"No het."* indicates the model without preference heterogeneity, *"Only δ"* the model with only heterogeneity in δ, *"Only γ"* the model with only heterogeneity in γ, *"No port."* the model without endogenous portfolio choice, *"No idio. ret."* the model without idiosyncratic returns and *"No skew."* the model without tail income shocks, stock market crashes and correlation between the income and return processes.

Understanding policies. In order to shed some light on the mechanisms driving the results for the model without preference heterogeneity, it is useful to examine the policy functions.

There are two well known factors (see, e.g., Campbell and Viceira, 2002) shaping the optimal choice of the risky share: the human capital-wealth ratio and the extent to which human capital has bond-like properties (i.e., should human capital be considered more similar to a safe or risky asset). This follows from the fact that optimal consumption is a function of such ratio, so its level and riskiness matter for consumption smoothing.

In a model where income is bond-like (see for example the benchmark specification in Cocco et al., 2005), the optimal risky share is 100% for the wealth-poor, and gradually declines as human capital gets smaller relative to wealth. The specification we use, however, features cyclically skewed income shocks, which helps to generate realistic portfolio choices as first shown in Catherine (2021). Intuitively, if the most adverse income shocks are more likely to happen in times of bad stock market performance, agents with low wealth must be more careful in their portfolio decisions.

To visualize this effect and to show that we replicate the findings of Catherine (2021), Figure 3.3 presents the policy functions for the risky share (ξ) over cash-on-hand for three different values - low, medium and high - of the persistent component of idiosyncratic income. For the sake of the argument, note that the x-axis is in log scale and that the participation cost is temporarily set to a very low value to make patterns more visible.

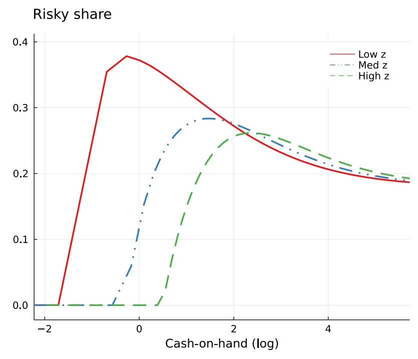


Figure 3.3: Policy functions for the risky share (ξ) for different states of the persistent component of idiosyncratic income. The three z values approximately correspond to the 25th, 50th and 75th percentiles of the steady state persistent income distribution.

In contrast to the bond-like human capital case, the optimal risky share for all three income states starts at zero, then gradually increases and finally decreases in cash-on-hand. The strongest effect is for agents with a higher z state and, for those among them with low levels of cash-on-hand, the standard finding of a positive relationship between human capital and the risky share is even reversed. This is because negatively skewed income shocks are especially severe for agents with higher persistent income as a fraction of total wealth.

There are two conclusions to draw from this discussion. Firstly, even lacking preference heterogeneity, the cyclical skewness channel helps to match portfolio patterns for the wealth-poor, as the extreme risk-taking of agents with a high human capital-to-wealth ratio implied by more traditional income processes is reduced. However, for most persistent income states, the increase in optimal risky share happens over a relatively narrow range of wealth (note the log scale in Figure 3.3). Thus - secondly - even if participation costs are absent, this mechanism alone has difficulties matching the gradual increase in the empirical risk-taking patterns over the whole wealth distribution. In particular, since any channel relying on the nature of human capital is by construction weak for agents with a low human capital-to-wealth ratio, without preference heterogeneity the model cannot match any increase at the top of the wealth distribution, which highlights the importance of this latter element for our results.

3.5.2 Fixed portfolio choice

To what extent does optimal portfolio choice amplify the effect of preference heterogeneity on wealth inequality? In this section, we study the counterfactual predictions of the benchmark model without endogenous portfolio choice. Specifically, we solve the model fixing for all individuals the share invested in the risky asset such that the ratio of risky assets held in the economy to total net wealth equals that of the benchmark specification.[22]

The impact on wealth inequality of optimal portfolio choice is significant, particularly at the top. Table 3.4 shows that under fixed portfolio choices, the wealth Gini decreases from 0.64 in the benchmark model to 0.58 and that the share of wealth held by the top quintile and by the top 1% decreases, respectively, from 67.5% to 62.5% and from 20.1% to 16.1%.[23]

The explanation for this finding is the following. Type-two individuals have a higher optimal risky share, which results in higher aver-

[22]Since all the targeted moments (except for the wealth-to-income ratio) are related to portfolio choice, we do not re-estimate the model parameters for this counterfactual, but use the same values of the benchmark case.

[23]Note that in an alternative counterfactual with the share of risky assets fixed at zero, the effect on wealth inequality would be even larger.

age returns than the rest of the population and amplifies the impact of larger saving rate due to lower impatience, as well as the large idiosyncratic shocks they are exposed to. Thus, when they are not forced to choose the same portfolio composition as the rest of the population, they are more likely to land on the top of the wealth distribution. The quantitative importance of these channels is illustrated in Figure 3.4, which shows that the concentration of type-two individuals among the wealthy is strongest in the benchmark model with endogenous portfolio choice.

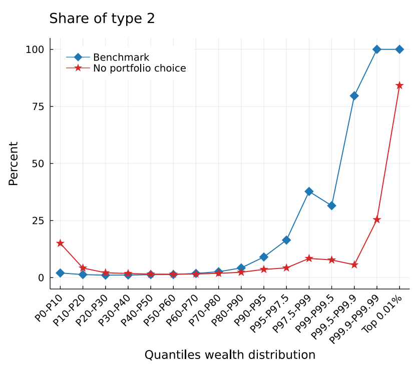


Figure 3.4: Share of type 2 individuals, benchmark model vs. model without portfolio choice.

As a result (see Figure 3.1), in line with the empirical patterns, in the model with preference heterogeneity and free portfolio choice, the average unconditional risky share is an increasing function of wealth, even at the very top of the distribution.

3.5.3 No idiosyncratic returns

To understand the role of idiosyncratic returns in shaping the wealth distribution, we evaluate the performance of another counterfactual model without idiosyncratic returns. Specifically, we set ζ equal to zero, which, in turn, implies zero mean and variance of log idiosyncratic returns $\eta_{i,t}$ for all individuals i and periods t.[24] As depicted in Figure 3.5, the model-generated policies match the data almost as accurately as the benchmark. The estimates of the common parameters across the two specifications are also similar, as reported in Table 3.3.

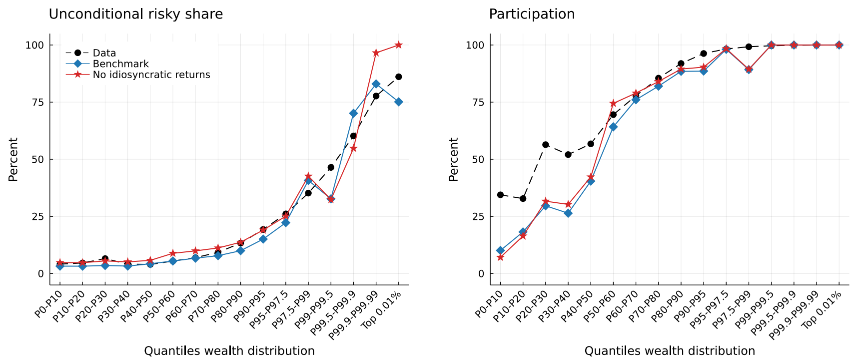


Figure 3.5: Policies over the wealth distribution in the model without idiosyncratic return risk, compared to the benchmark and data.

Table 3.4 shows the effect on the wealth distribution. As participation is low for the bottom quantiles in the benchmark and the counterfactual model, idiosyncratic returns barely affect the share of wealth held by the lowest three quintiles. The effect is largest for the top quantiles where agents participate more. For the top 1% of the wealth distribution, the share of wealth held declines from 20.1% in the benchmark to 18.3% in the model without idiosyncratic returns. When looking at total wealth inequality as measured by the Gini coefficient, the decrease in inequality is not too pronounced: 0.64 in the benchmark vs. 0.62 in the counterfactual. The limited impact of the idiosyncratic component of returns is in line with the results in Hubmer et al. (2021), who also find a small effect and mainly clustered at the top.

[24] As there is no idiosyncratic return risk, when we estimate this restricted model we only target the schedules of risky share and participation over the wealth distribution (and wealth-to-income ratio).

3.5.4 Heterogeneity in one preference parameter

To highlight the effects coming from impatience and risk aversion in isolation, we solve for counterfactuals where we restrict the model to heterogeneity in a single preference parameter (δ, γ). In particular, we re-estimate the model parameters two times and in each estimation we allow for heterogeneity in just one of them. We target the same moments as in the benchmark model.

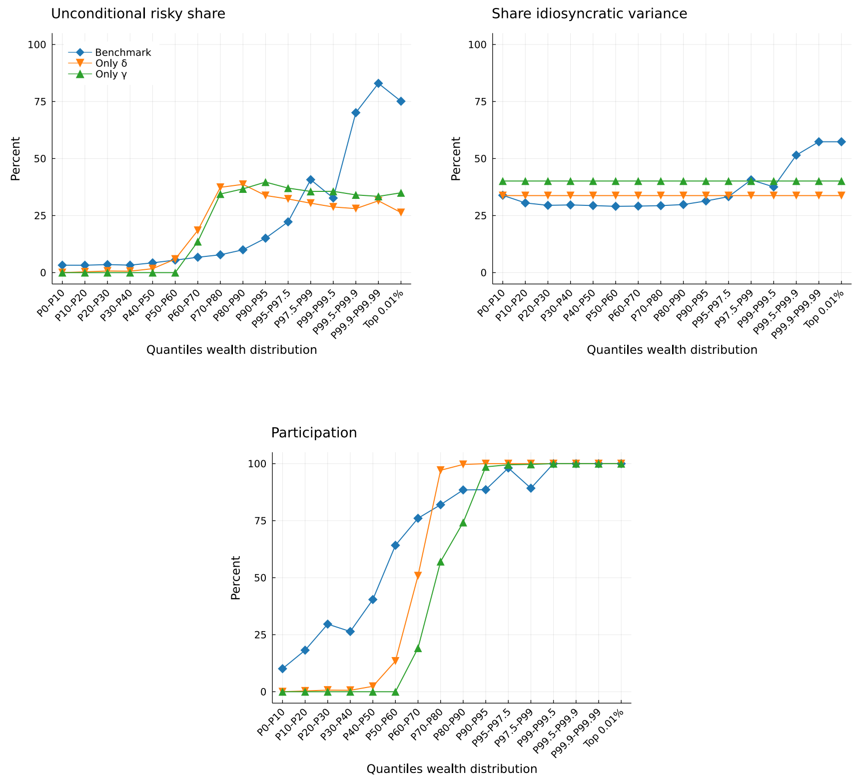


Figure 3.6: Policies over the wealth distribution. Specifications allowing for heterogeneity in one preference parameter at a time, compared to the benchmark.

The estimated parameters and the effect on the wealth distribution are presented in Tables 3.3 and 3.4 in the columns "Only δ" and "Only γ". We will analyze more in detail below the results from the two specifications, but it is already worth noticing that both these counterfactual models generate higher wealth inequality (especially the former) without providing a good fit to empirical portfolio choice patterns, as illustrated in Figure 3.6.

Only δ. A low discount factor implies both low saving rates and low participation, so in theory this setup can make the type more willing to invest into stocks concentrate on the top of the wealth distribution. However, since this margin does not affect substantially the conditional risky share, the increasing risky share pattern is matched only through the participation channel and hence it is not sufficiently gradual. Furthermore, after full participation the risky share cannot increase anymore in wealth. The attempt to match portfolio choice only through heterogeneity in the discount factor results in a rather extreme value for δ of the majority type, namely 0.75. Due to the resulting powerful separation between the two types over the wealth distribution, wealth inequality becomes large, even surpassing the empirical benchmark from Swedish data.[25]

Only γ. As high risk aversion implies low stock holdings, but high savings, generating a stock holder type on the top of the wealth distribution is less straightforward by heterogeneity in γ. Therefore portfolio choice patterns are matched again mostly through the participation margin, and the calibrated risk aversion parameters are estimated to match the average stock holdings over the region where participation occurs, achieving a slightly less pronounced decrease in risky share for the richest than in the "Only δ" case.

[25]That δ heterogeneity can generate high wealth inequality is known at least from Krusell and Smith (1998).

3.5.5 No skewness in labor income and return

Labor income and returns in the benchmark model are skewed. We have discussed the implications of these properties for the policy functions in section 3.5.1. In this section, we turn off skewness in labor income and returns and assess how well this restricted model performs. To be precise, we turn off stock market crashes, tail income shocks and the correlation between the income and return process.[26]

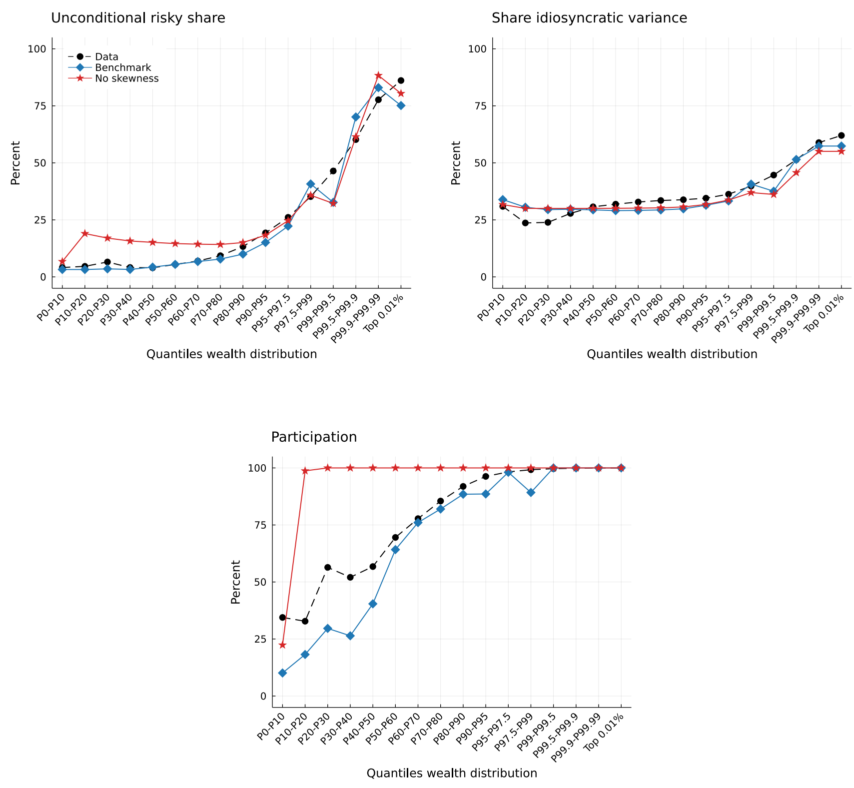


Figure 3.7: Policies over the wealth distribution when skewness and correlations are turned off, compared to the benchmark and data.

As for the previous counterfactuals, the parameter estimates are

[26]While turning off skewness, we ensure that returns and labor income have the same mean and variance as in the benchmark model. For a detailed outline see Appendix 3.D.

shown in Table 3.3. Intuitively, to match a given schedule of participation, risk aversion or the fixed portfolio cost need to increase as skewness risk disappears. This is exactly what the parameter estimates point to. Risk aversion of type-one individuals increases to roughly 18, while the fixed cost of participation increases roughly five times relative to the benchmark case. Figure 3.7 shows the fit of the targeted portfolio choice patterns. The share of idiosyncratic variance in total return variance is well matched, however the model overshoots the unconditional risky share and participation for most percentiles of the wealth distribution.

Even though preference heterogeneity still enables us to accurately match inequality (see Table 3.4) and all portfolio patterns for the top 10% of the wealth distribution, without skewness and correlation in income and return risk the model struggles to provide a good fit on the lower portion of the wealth distribution. In particular, in line with the predictions of standard portfolio choice models (as discussed in section 3.5.1), between the initial jump due to increasing participation and the final hike due to the compositional effect, the average risky share is a counterfactually decreasing function of wealth.

3.6 Conclusion

This paper introduces a macroeconomic angle to the recent empirical findings on portfolio choice by answering the following questions. First, which additional model ingredients to an otherwise standard incomplete-markets model suffice to generate portfolio choice characteristics consistent with the data? Second, how do those different model ingredients help to generate a realistic wealth distribution?

We include heterogeneity in individual preferences and a rich process that features cyclical skewness in earnings and idiosyncratic returns à la Catherine (2021) to a Bewley model with endogenous portfolio choice. We estimate the parameters governing preference heterogeneity to match the portfolio choice patterns documented in Bach et al. (2020). Heterogeneity in patience, risk aversion and the desire to diversify idiosyncratic return risk are three examples of preference heterogeneity that *jointly* generate realistic portfolio choice patterns over the wealth distribution. Alternative model specifications that abstract from preference heterogeneity, endogenous portfolio choice and non-normalities in the shocks' distributions worsen the fit of the portfolio choice patterns considerably. The combination of preference heterogeneity and endogenous portfolio choice further yields a close match of the wealth distribution, particularly at the top.

This paper attempts to connect the household finance literature on portfolio choice and the macroeconomics literature on wealth inequality. We find that a key element to do that is preference heterogeneity: one type of individuals is characterized by a relatively high risk aversion and a low time preference rate - commonly used in the household finance literature - whereas the other, much less numerous type of individuals features preference parameters commonly used in the macroeconomics literature. Interestingly, this result is an outcome of our parameter estimation rather than an exogenous assumption.

The model presented in this paper captures salient features of portfolio choice in the data endogenously and delivers quantitative predictions on the distribution of wealth. In ongoing work, we quantify dynamic effects of changes in the environment, e.g., aggregate return shocks or modifications in the tax schedule, on the distribution of wealth along the transition, accounting for the optimal response in individuals' portfolio choices.

References

Aguiar, M., Bils, M., and Boar, C. (2021). Who are the hand-to-mouth? *Working Paper*.

Bach, L., Calvet, L., and Sodini, P. (2018). From saving comes having? disentangling the impact of savings on inequality. *SSRN Electronic Journal*.

Bach, L., Calvet, L. E., and Sodini, P. (2020). Rich pickings? risk, return, and skill in household wealth. *American Economic Review*, 110(9):2703–47.

Benhabib, J., Bisin, A., and Luo, M. (2019). Wealth distribution and social mobility in the us: A quantitative approach. *American Economic Review*, 109(5):1623–47.

Benhabib, J., Bisin, A., and Zhu, S. (2011). The distribution of wealth and fiscal policy in economies with finitely lived agents. *Econometrica*, 79(1):123–157.

Benhabib, J., Bisin, A., and Zhu, S. (2015). The wealth distribution in bewley economies with capital income risk. *Journal of Economic Theory*, 159:489–515.

Calvet, L., Campbell, J., and Sodini, P. (2007). Down or out: Assessing the welfare costs of household investment mistakes. *Journal of Political Economy*, 115(5):707–747.

Calvet, L. E., Campbell, J. Y., Gomes, F. J., and Sodini, P. (2021). The cross-section of household preferences. *Working Paper*.

Campbell, J. and Viceira, L. (2002). *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors*. Oxford University Press.

Catherine, S. (2021). Countercyclical Labor Income Risk and Portfolio Choices over the Life Cycle. *The Review of Financial Studies*.

Catherine, S., Sodini, P., and Zhang, Y. (2021). Countercyclical income risk and portfolio choices: Evidence from sweden. *Working Paper*.

Chang, Y., Hong, J., Karabarbounis, M., and Wang, Y. (2018). Income Volatility and Portfolio Choices. 2018 Meeting Papers 412, Society for Economic Dynamics.

Cocco, J. F., Gomes, F. J., and Maenhout, P. J. (2005). Consumption and Portfolio Choice over the Life Cycle. *The Review of Financial Studies*, 18(2):491–533.

De Nardi, M. and Fella, G. (2017). Saving and wealth inequality. *Review of Economic Dynamics*, 26:280–300.

Fagereng, A., Guiso, L., Malacrino, D., and Pistaferri, L. (2020). Heterogeneity and persistence in returns to wealth. *Econometrica*, 88(1):115–170.

Fagereng, A., Guiso, L., and Pistaferri, L. (2017). Portfolio Choices, Firm Shocks, and Uninsurable Wage Risk. *The Review of Economic Studies*, 85(1):437–474.

Gabaix, X., Lasry, J.-M., Lions, P.-L., and Moll, B. (2016). The dynamics of inequality. *Econometrica*, 84(6):2071–2111.

Gomes, F. and Michaelides, A. (2005). Optimal life-cycle asset allocation: Understanding the empirical evidence. *The Journal of Finance*, 60(2):869–904.

Guvenen, F., Karahan, F., Ozkan, S., and Song, J. (2021). What do data on millions of u.s. workers reveal about lifecycle earnings dynamics? *Econometrica*, 89(5):2303–2339.

Guvenen, F., Ozkan, S., and Song, J. (2014). The nature of countercyclical income risk. *Journal of Political Economy*, 122(3):621–660.

Hubmer, J., Krusell, P., and Smith., A. A. (2021). Sources of us wealth inequality: Past, present, and future. *NBER Macroeconomics Annual*, 35:391–455.

Krueger, D., Mitman, K., and Perri, F. (2016). Chapter 11 - macroeconomics and household heterogeneity. In Taylor, J. B. and Uhlig, H., editors, *Handbook of Macroeconomics*, volume 2, pages 843–921. Elsevier.

Krusell, P. and Smith, Jr., A. A. (1998). Income and wealth heterogeneity in the macroeconomy. *Journal of Political Economy*, 106(5):867–896.

Merton, R. (1969). Lifetime portfolio selection under uncertainty: The continuous-time case. *The Review of Economics and Statistics*, 51:247–57.

Nirei, M. and Aoki, S. (2016). Pareto distribution of income in neoclassical growth models. *Review of Economic Dynamics*, 20:25–42.

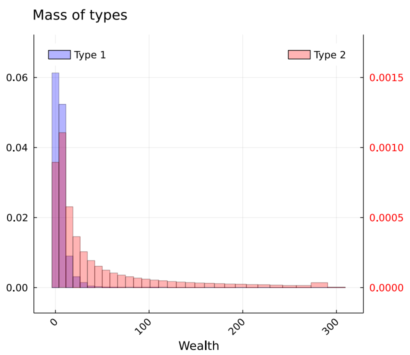
Vestman, R. (2018). Limited Stock Market Participation Among Renters and Homeowners. *The Review of Financial Studies*, 32(4):1494–1535.

Vissing-Jorgensen, A. (2003). Perspectives on behavioral finance: Does ”irrationality” disappear with wealth? evidence from expectations and actions. *NBER Macroeconomics Annual*, 18:139–194.

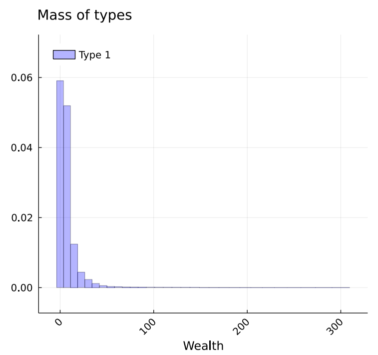
Appendices

3.A Additional figures

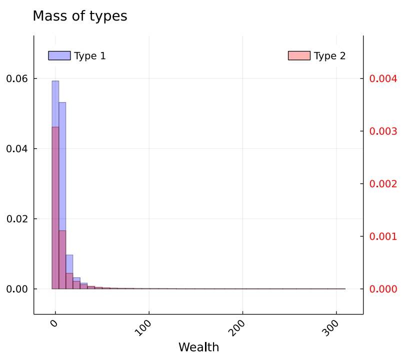
Panel (a): Benchmark



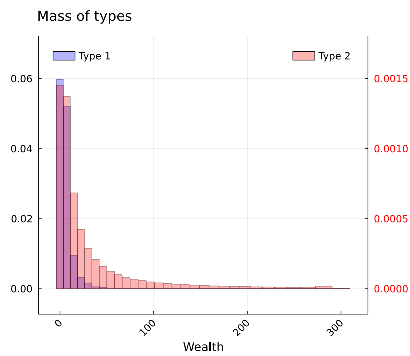
Panel (b): Homogeneous pref.



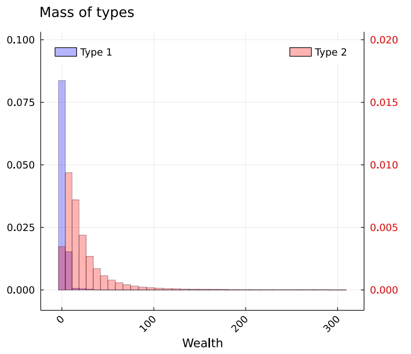
Panel (c): Fixed port. choice



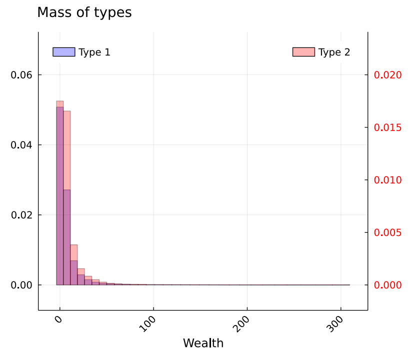
Panel (d): No idio. returns



Panel (e): Only δ heterogeneity



Panel (f): Only γ heterogeneity



Panel (g): No skewness

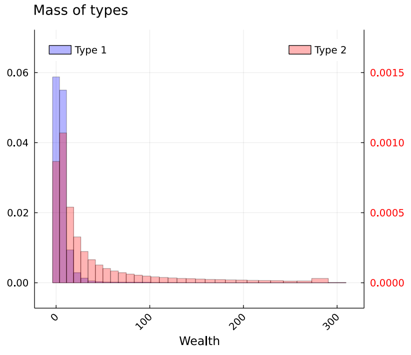


Figure 3.A.1: Mass of types over wealth (y-axis for Type 2 on the right) in different model specifications.

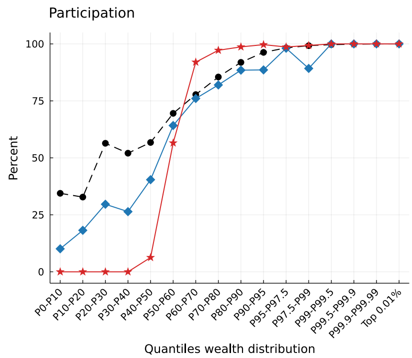


Figure 3.A.2: Participation over the wealth distribution in the model with homogeneous preferences (stars), compared to benchmark (diamonds) and data (circles).

3.B Numerical solution

Discretization and grids construction.

Normally distributed random variables. Let X be an i.i.d. Normally distributed random variable with mean μ_x and standard deviation σ_x. We discretize X using Gaussian quadrature. Specifically, the support of X is approximated with a finite grid of values $x_1, \ldots, x_{N_q}$ computed as follows:

$$x_j = \mu_x + \sqrt{2}\sigma_x Z_j, \quad j = 1, \ldots, N_q$$

where the Z_j's are Gauss-Hermite nodes and the probability mass of each point of the discretized support is computed as:

$$p(x_j) = \omega_j / \sqrt{\pi}, \quad j = 1, \ldots, N_q$$

where the ω_j's are Gauss-Hermite weights. This procedure is used to discretize $r_{2,t}$, ϕ_t, $\eta_{i,t}$ and the distributions of $\nu_{i,t}$ and $r_{1,t}$ conditional, respectively, on tail/non-tail event and on stock market crash/normal period. N_q is the same for all shocks.

Persistent component of idiosyncratic income. We approximate the process governing the evolution of $z_{i,t}$ as follows: (i) we discretize the conditional distribution of $\varepsilon_{i,t}$, (ii) we compute the evolution of the persistent component according to equation (3.5) and (iii) we evaluate the model functions at the resulting value of $z_{i,t}$. The advantage of this method is that it requires to discretize just the conditional distribution of $\varepsilon_{i,t}$, which is easier than discretizing the full process of $z_{i,t}$. In particular, the crucial connections between the higher moments of $z_{i,t}$ and other variables are preserved. The disadvantage is that the resulting values of $z_{i,t}$ will very often be off grid, so we need a grid of values that captures well the behavior of the model at such points given our interpolation procedure.[27]

Given the above discussion, the grids for $\varepsilon_{i,t}$ and $z_{i,t}$ are constructed as follows. The conditional distribution of $\varepsilon_{i,t}$ is discretized using the procedure described above for Normally distributed shocks with N_q points. To set up the grid for $z_{i,t}$, instead, we first construct an exponentially spaced grid of $(N_z - 1)/2 + 1$ points with minimum value equal to zero, maximum value equal to $z_{\max}$ and spacing parameter equal to spacing_z. This gives us the positive side of the grid plus the central point (which is therefore equal to zero). Then, we add the negative $(N_z - 1)/2$ values by taking the negative of the positive values just computed and obtain the full grid of N_z points.

Cash-on-hand and savings. For reasons that will be explained when describing the interpolation procedure, we need to keep track of the minimum value of cash-on-hand implied by each value in the grid of $z_{i,t}$. Thus, we construct N_z grids of cash-on-hand values - one for each grid value of $z_{i,t}$ - each of which is an exponentially spaced grid of N_a points with minimum value equal to the lowest possible realization of cash-on-hand - computed from equation (3.13) - implied by the specific grid value of $z_{i,t}$ under consideration, the borrowing

[27] See below for more details on the interpolation method.

limit and the discretized values of the shocks, maximum value equal to $a_{\max}$ and spacing parameter equal to spacing_a.

The grid for savings $s_{i,t}$ is an exponentially spaced grid of N_s points with minimum value equal to $\overline{s}$, maximum value equal to $s_{\max}$ and spacing parameter equal to spacing_s.

Table 3.B.1 summarizes our choices for the numerical parameters.

N_q	N_z	N_a	N_s	$z_{\max}$	$a_{\max}$	$s_{\max}$	spacing_z	spacing_a	spacing_s
5	15	350	N_a	3.5	300.0	300.0	1.6	1.25	1.25

Table 3.B.1: Numerical parameters.

Solving the optimization problem. Whenever it does not lead to confusion, we are dropping time and individual specific indices. To ease up exposition, we also drop the dependence of the value and policy functions (and thus of the preference parameters) on θ. First of all note that as w_t follows a random walk and utility is homogeneous, we can scale with the wage level and reduce the state-space by one dimension. Let us define $\hat{x} = x/\exp(w)$ for a generic variable x representing c, a, s and equivalently for log income, y, we define $\exp(\hat{y}) = \exp(y)/\exp(w)$. Also define

$$\widehat{V}(a, z) = V(a, z, 0)$$

so that we can write

$$V(a, z, w) = \exp(w)V\Big(\frac{a}{\exp(w)}, z, 0\Big) = \exp(w)\widehat{V}(\hat{a}, z).$$

Now the optimization problem can be written as

$$\widehat{V}(\hat{a}, z) = \max_{\{\hat{c}, \hat{s}, \xi\}} \left\{(1-\delta)\hat{c}^{1-\psi} + \delta\Big[\mathbb{E}\big[e^{(w'-w)(1-\gamma)}\widehat{V}(\hat{a}', z')^{1-\gamma}\big]\Big]^{\frac{1-\psi}{1-\gamma}}\right\}^{\frac{1}{1-\psi}}$$

subject to

$$
\begin{aligned}
\hat{c} + \hat{s} + F_{i,t}\hat{f} &= \hat{a} \\
\hat{a}' &= \Big[R^f + \xi R^{e\prime}\Big]\hat{s}e^{w-w'} + \exp(\hat{y}') + \hat{b}' \\
\hat{s} &\geq \bar{s}.
\end{aligned}
$$

To simplify ideas and notation, let us introduce

$$
\widetilde{V}(\hat{s}, \xi, z) = \Big[\mathbb{E}\big[e^{(w'-w)(1-\gamma)}\widehat{V}(\hat{a}', z')^{1-\gamma}\big]\Big]^{\frac{1-\psi}{1-\gamma}}
$$

The first order condition with respect to the risky share is then:

$$
\begin{aligned}
0 =& \frac{\partial \widetilde{V}(\hat{s}, \xi, z)}{\partial \xi} = \frac{1-\psi}{1-\gamma}\Big[\widetilde{V}(\hat{s}, \xi, z)\Big]^{\frac{\gamma-\psi}{1-\psi}} \\
& \times \mathbb{E}\Big[(1-\gamma)e^{(w'-w)(1-\gamma)}\widehat{V}(\hat{a}', z')^{-\gamma}\frac{\partial \widehat{V}(\hat{a}', z)}{\partial \hat{a}'}\frac{d\hat{a}'}{d\xi}\Big] \\
0 =& \mathbb{E}\Big[e^{-\gamma(w'-w)}\widehat{V}(\hat{a}', z')^{-\gamma}\frac{\partial \widehat{V}(\hat{a}', z)}{\partial \hat{a}'}R^{e\prime}\Big],
\end{aligned}
$$

where for the last equation we used that $\widetilde{V}(\hat{s}, \xi, z) \neq 0$. The first order condition for the consumption/saving decision reads:

$$
\begin{aligned}
(1-\delta)(1-\psi)\hat{c}^{-\psi} &= \delta\frac{\partial \widetilde{V}(\hat{s}, \xi, z)}{\partial \hat{s}} \\
(1-\delta)(1-\psi)\hat{c}^{-\psi} &= \delta\frac{1-\psi}{1-\gamma}\Big[\widetilde{V}(\hat{s}, \xi, z)\Big]^{\frac{\gamma-\psi}{1-\psi}} \\
&\quad \times \mathbb{E}\Big[(1-\gamma)e^{(w'-w)(1-\gamma)}\widehat{V}(\hat{a}', z')^{-\gamma}\frac{\partial \widehat{V}(\hat{a}', z)}{\partial \hat{a}'}\frac{d\hat{a}'}{d\hat{s}}\Big] \\
(1-\delta)\hat{c}^{-\psi} &= \delta\Big[\widetilde{V}(\hat{s}, \xi, z)\Big]^{\frac{\gamma-\psi}{1-\psi}} \\
&\quad \times \mathbb{E}\Big[e^{-\gamma(w'-w)}\widehat{V}(\hat{a}', z')^{-\gamma}\frac{\partial \widehat{V}(\hat{a}', z)}{\partial \hat{a}'}(R^f + \xi R^{e\prime})\Big]
\end{aligned}
$$

and the envelope condition is:

$$\frac{\partial \widehat{V}(\hat{a},z)}{\partial \hat{a}} = \frac{1}{1-\psi}\big[\widehat{V}(\hat{a},z)\big]^{\psi} \times \left[(1-\delta)(1-\psi)\hat{c}^{-\psi}\frac{d\hat{c}}{d\hat{a}} + \delta\Big[\frac{\partial \widetilde{V}(\hat{s},\xi,z)}{\partial \hat{s}}\frac{d\hat{s}}{d\hat{a}} + \frac{\partial \widetilde{V}(\hat{s},\xi,z)}{\partial \xi}\frac{d\xi}{d\hat{a}}\Big]\right]$$

$$\frac{\partial \widehat{V}(\hat{a},z)}{\partial \hat{a}} = (1-\delta)\big[\widehat{V}(\hat{a},z)\big]^{\psi}\hat{c}(\hat{a},z)^{-\psi}$$

After simplifying the two first order conditions read

$$0 = \mathbb{E}\Big[e^{-\gamma(w'-w)}\widehat{V}(\hat{a}',z')^{\psi-\gamma}\hat{c}'(\hat{a}',z')^{-\psi}R^{e\prime}\Big] \tag{3.B.1}$$

$$\hat{c}^{-\psi} = \delta\Big[\widetilde{V}(\hat{s},\xi,z)\Big]^{\frac{\gamma-\psi}{1-\psi}}\mathbb{E}\Big[e^{-\gamma(w'-w)}\widehat{V}(\hat{a}',z')^{\psi-\gamma}\hat{c}'(\hat{a}',z')^{-\psi}(R^{f}+\xi R^{e\prime})\Big] \tag{3.B.2}$$

Algorithm to solve for value and policy functions. To solve equations (3.B.1) and (3.B.2) we need to evaluate expectations over policy and value function both on and off the grid points for cash-on-hand and persistent income. The sections below provide further details on how we compute expectations and interpolate.

1. Assume we have a guess for $\widehat{V}$, $\hat{c}$ and ξ. For starting one can simply take $\hat{c}'(\hat{a},z) = \hat{a}$, $\widehat{V}'(\hat{a},z) = (1-\delta)^{\frac{1}{1-\psi}}\hat{a}$ with an arbitrary ξ' function. Fix a grid $\{\theta_1,\ldots,\theta_k,\ldots,\theta_K\}$ for the preference states, $\{z_1,\ldots,z_j,\ldots,z_M\}$ for the possible values of persistent income and $\{\hat{s}_1 = \bar{s}, \hat{s}_2,\ldots,\hat{s}_i,\ldots,\hat{s}_N\}$ for savings.

2. For all i, j (i.e., for any preference and persistent income state)

    (a) For all i (savings values) compute

        i. the optimal risky share ξ (under participation only, for non-participation set $\xi = 0$ and go to ii.). Recall that ξ is chosen to maximize $\widetilde{V}(\hat{s},\xi,z)$ and that (ignoring constants)

$$\frac{\partial \widetilde{V}(\hat{s}, \xi, z)}{\partial \xi} = \mathbb{E}\Big[e^{-\gamma(w'-w)}\widehat{V}(\hat{a}', z')^{\psi-\gamma}\hat{c}'(\hat{a}', z')^{-\psi}R^{e\prime}\Big].$$

From the second order condition it follows that there is a unique local maximum. Optimal risky share is computed as follows:

A. if a risky share of 1 was optimal in the previous iteration, check whether this is still true. This is the case if

$$\frac{\partial \widetilde{V}(\hat{s}, \xi, z)}{\partial \xi} > 0.$$

If not, save the information that $\xi < 1$.

B. if a risky share of 0 was optimal in the previous iteration, check whether this is still true. This is the case if

$$\frac{\partial \widetilde{V}(\hat{s}, \xi, z)}{\partial \xi} < 0.$$

If not, save the information that $\xi > 0$.

C. if in the previous iteration neither $\xi = 0$ nor $\xi = 1$ was optimal, use the secant method (combined with the information from i. and ii.) to find ξ such that

$$\frac{\partial \widetilde{V}(\hat{s}, \xi, z)}{\partial \xi} = 0.$$

For the secant method two starting points are needed. As the first point use the previous iteration's optimal risky share. The second point is found by moving slightly to the left or right of the first point (depending on the sign of $\frac{\partial \widetilde{V}(\hat{s},\xi,z)}{\partial \xi}$).

ii. optimal consumption $\hat{c}$ by solving

$$\hat{c}^{-\psi} = \delta\Big[\widetilde{V}(\hat{s},\xi,z)\Big]^{\frac{\gamma-\psi}{1-\psi}} \times \mathbb{E}\Big[e^{-\gamma(w'-w)}\widehat{V}(\hat{a}',z')^{\psi-\gamma}\hat{c}'(\hat{a}',z')^{-\psi}(R^f+\xi R^{e\prime})\Big]$$

With this at hand we can compute the value function as

$$\widehat{V} = \left\{(1-\delta)\hat{c}^{1-\psi} + \delta\Big[\mathbb{E}\big[e^{(w'-w)(1-\gamma)}\widehat{V}'(\hat{a}',z')^{1-\gamma}\big]\Big]^{\frac{1-\psi}{1-\gamma}}\right\}^{\frac{1}{1-\psi}}$$

and cash-on-hand as

$$\hat{a} = \hat{c} + \hat{s} + \hat{F}f$$

Note that for computing cash-on-hand the budget constraint for participation and non-participation differ wrt. to the participation costs. Due to the discrete choice of whether to participate or not, computation of consumption, value and cash-on-hand has to be done separately assuming using the optimal risky share computed in (i) and assuming non-participation.

(b) Interpolate the value function and the policy functions for consumption and the risky share (separately for participation and non-participation) over cash-on-hand. By comparing value functions over the cash-on-hand grid we obtain at which part of the grid the agent is participating. We connect value functions and policy functions over the cash-on-hand grid at an ϵ environment around the participation threshold.

3. Convergence is declared when the absolute change in the optimal risky share and the relative change of optimal consumption policies are both smaller than a pre-specified tolerance at every grid point.

Changes when $\psi = 1$. The value function satisfies:

$$\widehat{V}(\hat{a}, z) = \max_{\{\hat{c},\hat{s},\xi\}} \left\{ \hat{c}^{1-\delta} \Big[\mathbb{E}\big[e^{(w'-w)(1-\gamma)} \widehat{V}(\hat{a}', z')^{1-\gamma} \big] \Big]^{\frac{\delta}{1-\gamma}} \right\}$$

$$\widetilde{V}(\hat{s}, \xi, z) = \Big[\mathbb{E}\big[e^{(w'-w)(1-\gamma)} \widehat{V}(\hat{a}', z')^{1-\gamma} \big] \Big]^{\frac{\delta}{1-\gamma}}$$

The first order condition with respect to the risky share is:

$$\begin{aligned}
0 = & \frac{\partial \widetilde{V}(\hat{s}, \xi, z)}{\partial \xi} = \hat{c}^{1-\delta} \frac{\delta}{1-\gamma} \Big[\widetilde{V}(\hat{s}, \xi, z) \Big]^{\frac{\gamma-1+\delta}{\delta}} \\
& \times \mathbb{E}\Big[(1-\gamma) e^{(w'-w)(1-\gamma)} \widehat{V}(\hat{a}', z')^{-\gamma} \frac{\partial \widehat{V}(\hat{a}', z')}{\partial \hat{a}'} \frac{d\hat{a}'}{d\xi} \Big] \\
0 = & \mathbb{E}\Big[e^{-\gamma(w'-w)} \widehat{V}(\hat{a}', z')^{-\gamma} \frac{\partial \widehat{V}(\hat{a}', z')}{\partial \hat{a}'} R^{e\prime} \Big],
\end{aligned}$$

where for the last equation we used that $\widetilde{V}(\hat{s}, \xi, z) \neq 0$. The first order condition for the consumption/saving decision reads:

$$\begin{aligned}
(1-\delta)\hat{c}^{-\delta} \widetilde{V}(\hat{s}, \xi, z) &= \hat{c}^{1-\delta} \frac{\partial \widetilde{V}(\hat{s}, \xi, z)}{\partial \hat{s}} \\
(1-\delta)\hat{c}^{-1} \widetilde{V}(\hat{s}, \xi, z) &= \frac{\delta}{1-\gamma} \Big[\widetilde{V}(\hat{s}, \xi, z) \Big]^{\frac{\gamma-1+\delta}{\delta}} \\
& \qquad \times \mathbb{E}\Big[(1-\gamma) e^{(w'-w)(1-\gamma)} \widehat{V}(\hat{a}', z')^{-\gamma} \frac{\partial \widehat{V}(\hat{a}', z')}{\partial \hat{a}'} \frac{d\hat{a}'}{d\hat{s}} \Big] \\
(1-\delta)\hat{c}^{-1} &= \delta \Big[\widetilde{V}(\hat{s}, \xi, z) \Big]^{\frac{\gamma-1}{\delta}} \\
& \qquad \times \mathbb{E}\Big[e^{-\gamma(w'-w)} \widehat{V}(\hat{a}', z')^{-\gamma} \frac{\partial \widehat{V}(\hat{a}', z')}{\partial \hat{a}'} (R^f + \xi R^{e\prime}) \Big]
\end{aligned}$$

and the envelope condition is:

$$\begin{aligned}
\frac{\partial \widehat{V}(\hat{a}, z)}{\partial \hat{a}} &= (1-\delta)\hat{c}^{-\delta} \widetilde{V}(\hat{s}, \xi, z) \frac{d\hat{c}}{d\hat{a}} + \hat{c}^{1-\delta} \Big[\frac{\partial \widetilde{V}(\hat{s}, \xi, z)}{\partial \hat{s}} \frac{d\hat{s}}{d\hat{a}} + \frac{\partial \widetilde{V}(\hat{s}, \xi, z)}{\partial \xi} \frac{d\xi}{d\hat{a}} \Big] \\
\frac{\partial \widehat{V}(\hat{a}, z)}{\partial \hat{a}} &= (1-\delta)\hat{c}^{-\delta} \widetilde{V}(\hat{s}, \xi, z)
\end{aligned}$$

After simplifying the two first order conditions read:

$$0 = \mathbb{E}\Big[e^{-\gamma(w'-w)}\widehat{V}(\hat{a}',z')^{1-\gamma}\hat{c}'(\hat{a}',z')^{-1}R^{e\prime}\Big] \tag{3.B.3}$$

$$\hat{c}^{-1} = \delta\Big[\widetilde{V}(\hat{s},\xi,z)\Big]^{\frac{\gamma-1}{\delta}}\mathbb{E}\Big[e^{-\gamma(w'-w)}\widehat{V}(\hat{a}',z')^{1-\gamma}\hat{c}'(\hat{a}',z')^{-1}(R^f+\xi R^{e\prime})\Big] \tag{3.B.4}$$

Interpolation. The solution procedure outlined below will very often require to evaluate the value function and the consumption policy at points off the grid. As explained above, we do not discretize the persistent component of idiosyncratic income, which implies that we need to interpolate these functions not only at points off the cash-on-hand grid, but also off the grid of persistent income. In other words, we need a 2-dimensional interpolation procedure over the (a, z) grid. This is achieved by 2-dimensional linear interpolation.

Computing expectations. In order to solve the model, it is necessary to compute expectations of some non trivial functions. In the most general case, we need to compute expectations with respect the shocks r_1, r_2, ϕ, ε, ν and η.[28] To do that, we proceed as follows: (i) for all the possible combinations of grid values of these variables, we compute the value of the function (ii) we multiply it by the probability of that particular combination of values (iii) once we have done this for all the possible combinations we sum up all the function values obtained. Note that except for transitions in the preference state, the grid values and probabilities of the other shocks coincide with Gaussian quadrature nodes and weights, which enables us to compute expectations very accurately. Finally, remember that the distributions of r_1 is conditional on the realization or not of a stock market crash and, similarly, those of ε and ν on the realization of a tail event or not. This is taken into account simply

[28]Cases in which we do not need to take expectations with respect to one or more of these variables can be handled by the same procedure outlined here with straightforward modifications.

by scaling the probability of the discretized conditional distributions of these variables by the probability of these events.

Simulation and stationary distributions. In this model the unique aggregate state is the distribution of agents across individual states which are characterized by the triple $(\theta_{i,t}, z_{i,t}, a_{i,t})$. As described before, values of all these states are approximated by a finite grid, therefore in the numerical setting the distribution object can be described as a vector of length $N_t \cdot N_z \cdot N_a$ containing the probability weights corresponding to each individual state. To simulate the economy we need to compute the transition probabilities of moving from one individual state to another, which naturally depend on the actual value of the aggregate shocks. Therefore to examine the dynamic properties of the distribution and to characterize the quasi steady state distribution we need to construct transition matrices corresponding to all values of the aggregate shocks we want to simulate and then aggregate them into a quasi steady state transition matrix.

Conditional transition matrices Taken a realization of the aggregate shocks (r_1, r_2, ϕ) given for each individual state we can compute future cash-on-hand and future persistent income corresponding to each realization of idiosyncratic shocks, where these are simulated from the grids described above. Since in the generic case the simulated cash-on-hand and z values do not fall on grid point, the conditional probability weight corresponding to each simulated (a, z) pair is distributed between the neighboring 4 (or on edges of the grid 2) points proportionally to their relative distance. To avoid extrapolation errors, z is truncated between $-z_{max}$ and z_{max}. From the transition probabilities of moving from one individual state to another we can build up the transition matrices conditional on any realisation of the aggregate shocks.

Unconditional transition matrix To obtain a steady state distribution we need an "average transition matrix". One way of defining one would be taking the conditional transition matrix corresponding to the average values of all aggregate shocks. However, a steady state computed from such a matrix would miss all the consequences of cyclical movements in moments of idiosyncratic shocks, central to our analysis. Therefore we compute our steady state matrix as a weighted average of the conditional transition matrices corresponding to shock values used in the policy iteration, where weights are the probabilities that the given combination of shocks takes place. Hence all entries in the steady state matrix are the true unconditional transition probabilities of moving from one individual state to another (before knowing the shock values).

Steady state distribution The steady state is found by iteration, i.e., multiplying an arbitrary vector with the unconditional transition matrix until convergence. Note that the aggregate state object is a distribution over preference type, persistent income and cash-on-hand. Using the optimal saving policy function, we can compute steady state distribution over preference type, persistent income and end-of-period assets, which is what we refer to as wealth distribution.

3.C Distribution of total return

Denote the pdf of a normal distribution with mean μ and variance σ^2 with $f_{N(\mu,\sigma^2)}$ and the pdfs of r_1, r_2 and η with f_{r_1}, f_{r_2} and f_η. Since the latter three random variables are independent, we can write the joint pdf of (r_1, r_2, η) as:

$$\begin{aligned} f_{(r_1,r_2,\eta)}(r_1, r_2, \eta) =& f_{r_1}(r_1) f_{r_2}(r_2) f_\eta(\eta) = \\ =& p_r f_{N(\underline{\mu}_r,\sigma_{r1}^2)} f_{N(0,\sigma_{r2}^2)} f_{N(-\zeta^2\sigma_r^2/2,\zeta^2\sigma_r^2)} + \\ &+ (1-p_r) f_{N(\overline{\mu}_r,\sigma_{r1}^2)} f_{N(0,\sigma_{r2}^2)} f_{N(-\zeta^2\sigma_r^2/2,\zeta^2\sigma_r^2)} \end{aligned}$$

where $\sigma_r^2 = \sigma_{r_1}^2 + p_r\underline{\mu}_r^2 + (1-p_r)\overline{\mu}_r^2 - \mu_r^2 + \sigma_{r_2}^2$ and $\mu_r = p_r\underline{\mu}_r + (1-p_r)\overline{\mu}_r$. By properties of normal distirbutions, it follows that $r_1 + r_2 + \eta$ follows a mixed normal distribution with pdf:

$$\begin{aligned} f_{(r_1+r_2+\eta)}(r) = {} & p_r f_{N(\underline{\mu}_r - \zeta^2\sigma_r^2/2, \sigma_{r1}^2 + \sigma_{r2}^2 + \zeta^2\sigma_r^2)}(r) \\ & + (1-p_r) f_{N(\overline{\mu}_r - \zeta^2\sigma_r^2/2, \sigma_{r1}^2 + \sigma_{r2}^2 + \zeta^2\sigma_r^2)}(r) \end{aligned}$$

and total risky return $R_r = \exp(r_1 + r_2 + \eta)$ follows a corresponding mixed log-normal distribution. Therefore:

$$\begin{aligned} \mathbb{E}\left[R_r\right] = {} & p_r \exp\left(\underline{\mu}_r - \zeta^2\sigma_r^2/2 + (\sigma_{r1}^2 + \sigma_{r2}^2 + \zeta^2\sigma_r^2)/2\right) + \\ & + (1-p_r)\exp\left(\overline{\mu}_r - \zeta^2\sigma_r^2/2 + (\sigma_{r1}^2 + \sigma_{r2}^2 + \zeta^2\sigma_r^2)/2\right) = \\ = {} & p_r \exp\left(\underline{\mu}_r + (\sigma_{r1}^2 + \sigma_{r2}^2)/2\right) + (1-p_r)\exp\left(\overline{\mu}_r + (\sigma_{r1}^2 + \sigma_{r2}^2)/2\right) \end{aligned}$$

and

$$\begin{aligned} \mathbb{E}\left[R_r^2\right] = {} & p_r \exp\left(2\underline{\mu}_r - \zeta^2\sigma_r^2 + 2(\sigma_{r1}^2 + \sigma_{r2}^2 + \zeta^2\sigma_r^2)\right) + \\ & + (1-p_r)\exp\left(2\overline{\mu}_r - \zeta^2\sigma_r^2 + 2(\sigma_{r1}^2 + \sigma_{r2}^2 + \zeta^2\sigma_r^2)\right) = \\ = {} & p_r \exp\left(2\underline{\mu}_r + 2\sigma_{r1}^2 + 2\sigma_{r2}^2 + \zeta^2\sigma_r^2\right) \\ & + (1-p_r)\exp\left(2\overline{\mu}_r + 2\sigma_{r1}^2 + 2\sigma_{r2}^2 + \zeta^2\sigma_r^2\right) \end{aligned}$$

so that:

$$\begin{aligned} \mathbb{V}\left[R_r\right] = {} & \mathbb{E}\left[R_r^2\right] - \left(\mathbb{E}\left[R_r\right]\right)^2 = p_r \exp\left(2\underline{\mu}_r + 2\sigma_{r1}^2 + 2\sigma_{r2}^2 + \zeta^2\sigma_r^2\right) + \\ & + (1-p_r)\exp\left(2\overline{\mu}_r + 2\sigma_{r1}^2 + 2\sigma_{r2}^2 + \zeta^2\sigma_r^2\right) - \\ & - p_r^2 \exp\left(2\underline{\mu}_r + \sigma_{r1}^2 + \sigma_{r2}^2\right) - (1-p_r)^2\exp\left(2\overline{\mu}_r + \sigma_{r1}^2 + \sigma_{r2}^2\right) - \\ & - 2p_r(1-p_r)\exp\left(\underline{\mu}_r + \overline{\mu}_r + \sigma_{r1}^2 + \sigma_{r2}^2\right) \end{aligned}$$

Assuming $p_r = 1$ we would have:

$$\begin{aligned}
\mathbb{V}\left[R_r\right] &= \exp\left(2\mu_r + \sigma_{r1}^2 + \sigma_{r2}^2\right)\left(\exp\left(\sigma_{r1}^2 + \sigma_{r2}^2 + \zeta^2\sigma_r^2\right) - 1\right) \approx \\
&\approx \exp\left(2\mu_r + \sigma_{r1}^2 + \sigma_{r2}^2\right)\left(\sigma_{r1}^2 + \sigma_{r2}^2 + \zeta^2\sigma_r^2\right) = \\
&= \left(1 + \zeta^2\right)\exp\left(2\mu_r + \sigma_r^2\right)\sigma_r^2
\end{aligned}$$

which is intuitive.

3.D Counterfactuals

In this appendix we describe how we run the counterfactual experiments listed in section 3.5.

No idiosyncratic returns. To shut down idiosyncratic return shocks, we simply set $\zeta(\theta_{i,t}) = 0$ for all i and t.

No skewness in labor income and return.

Stock market crashes. To eliminate stock market crashes, the distribution of r_1 should be non-skewed while retaining its mean and variance from the benchmark specification with no preference heterogeneity. This is done by setting:

$$r_{1,t} \overset{i.i.d.}{\sim} \mathcal{N}\left(\mu_r, \tilde{\sigma}_{r_1}^2\right)$$

where $\mu_r = p_r\underline{\mu}_r + (1-p_r)\overline{\mu}_r$ and $\tilde{\sigma}_{r_1}^2 = \sigma_{r_1}^2 + p_r\underline{\mu}_r^2 + (1-p_r)\overline{\mu}_r^2 - \mu_r^2$.

Connection between wages and returns. Shutting down the connection between wages and returns implies that equation (2.6) should be replaced with:

$$w_t = g + w_{t-1} + \lambda_{rw}\mu_r + \phi_t$$

Skewness in idiosyncratic shocks. To turn off the skewness of idiosyncratic shocks, the distributions of ε and ν and should be replaced by non-skewed distributions having identical first and second moments as in the benchmark parametrization with no preference heterogeneity. Therefore:

$$\varepsilon_{i,t} \overset{i.i.d.}{\sim} \mathcal{N}\left(0, \sigma_\varepsilon^2\right) \qquad \nu_{i,t} \overset{i.i.d.}{\sim} \mathcal{N}\left(0, \sigma_\nu^2\right)$$

where

$$\sigma_\nu^2 = p_\varepsilon \underline{\sigma}_\nu^2 + (1 - p_\varepsilon)\overline{\sigma}_\nu^2$$

and

$$\begin{aligned}
\sigma_\varepsilon^2 &= p_\varepsilon \underline{\sigma}_\varepsilon^2 \\
&\quad + (1 - p_\varepsilon)\overline{\sigma}_\varepsilon^2 + p_\varepsilon(\mu_\varepsilon + \lambda_{\varepsilon w}(g + \lambda_{rw}\mu_r))^2 \\
&\quad + (1 - p_\varepsilon)\Big(- \frac{p_\varepsilon}{1 - p_\varepsilon}(\mu_\varepsilon + \lambda_{\varepsilon w}(g + \lambda_{rw}\mu_r))\Big)^2 \\
&= p_\varepsilon \underline{\sigma}_\varepsilon^2 + (1 - p_\varepsilon)\overline{\sigma}_\varepsilon^2 + \frac{p_\varepsilon}{1 - p_\varepsilon}(\mu_\varepsilon + \lambda_{\varepsilon w}(g + \lambda_{rw}\mu_r))^2
\end{aligned}$$

Sammanfattning

Den här avhandlingen består av tre fristående uppsatser. Samtliga kombinerar användningen av mikro- och makrodata och kvantitativa modeller för att studera hur aktörer balanserar sin utsatthet för inkomstrisk när de står inför idiosynkratiska och aggregerade chocker.

Det första kapitlet *"Effekterna på inkomsterna skiljer sig beroende på konjunkturläget" (Business cycle asymmetry of earnings pass-through)* analyserar hur endogena inkomstrisker uppstår ur den optimala fördelningen av riskdelning mellan arbetstagare och företag.

Medan förståelsen för hur mycket företag försäkrar sina arbetstagare mot lönefluktuationer är en fråga som existerat länge inom nationalekonomi, studerar denna artikel denna fråga ur en ny infallsvinkel genom att studera hur företags förmåga att göra detta varierar över konjunkturcykeln. Eftersom arbetsmarknads- och finansiella friktioner är bindande på olika sätt i hög- och lågkonjunkturer finns det skäl att misstänka att överföringen av chockerna till arbetstagarnas inkomster kan variera med de aggregerade chockerna.

Genom att använda svenska administrativa data dokumenterar jag att effekterna av idiosynkratiska företagschocker på arbetstagarnas inkomster faktiskt är asymmetriska över konjunkturcykeln. Företagen försäkrar arbetstagarna mot negativa chocker i perioder av icke-lågkonjunkturer men de gör detta i mycket mindre utsträckning vid konjunkturnedgångar. Positiva chocker, å andra sidan, delas speciellt med anställda om de är avsevärda, och detta gäller oavsett samhällsekonomins tillstånd.

Jag visar vidare att dessa empiriska mönster kan förklaras med en riktad sökmodell av arbetsmarknaden med jobbsökande av de som redan har ett jobb, riskaverta arbetstagare och företagens åtaganden. Nyckelelementet i modellen är den avvägning som företagen står inför när de väljer villkoren för anställningsförhållandet med sina arbetstagare. Å ena sidan, att försäkra riskaverta arbetstagare mot lönefluktuationer gör det möjligt att undvika att betala den volatilitetsriskpremie som impliceras av urholkningen i nyttofunktionen. Å andra sidan, att göra detta gör det inte möjligt att justera arbetstagarnas arbetsincitament i enlighet med företagens önskan att upprätthålla matchningen eller inte.

Utöver att matcha de asymmetrier i inkomsteffekterna som vi funnit i data, ger modellen en ny förklaring till konjunkturcykeltrenderna i inkomstrisk som dokumenterats nyligen i den empiriska litteraturen. Mer i detalj, eftersom effekterna av negativa chocker i genomsnitt är högre i konjunkturnedgångar och effekterna av positiva chocker är acykliska, visar den modellgenererade fördelningen av löneökningen en procyklisk skevhet. Slutligen värderar jag välfärdskostnaderna av konjunkturcykler och finner att de är avsevärda inom detta ramverk.

Det andra kapitlet, *"Att härleda inkomstegenskaper utifrån portföljval" (Inferring income properties from portfolio choices)*, visar att endogena inkomstrisker till följd av aktörers portföljval ger information om arbetsinkomstprocessens verkliga karaktär.

Även om den litteratur som försöker förstå inkomstprocessens karaktär är omfattande så har två huvudsakliga hypoteser uppstått: enligt den ena är inkomstchocker mycket beständiga och aktörerna har liknande livscykelprofiler – begränsade inkomstprofiler (restricted income profiles RIP); enligt den andra är inkomstschocker inte särskilt beständiga och livscykelprofilerna är individspecifika – heterogena inkomstprofiler (heterogeneous income profiles HIP).

I denna uppsats studerar jag huruvida aktörers portföljval innehåller relevant information för att bedöma vilket av de två synsätten som har mest stöd i data. Huvudidén är att eftersom olika typer av inkomstrisk innebär olika portföljallokeringsbeslut, kan forskaren sluta sig till inkomstprocessens karaktär genom att titta på det senare.

Efter att ha tagit den cykliska skevheten i beaktande i framtagandet av chockernas beständighet för HIP och RIP finner jag att profilerna för medianen av och variansen i konsumtionen över livscykeln liknar varandra mycket och sålunda inte har en stark identifikationskraft.

Emellertid så innebär HIP och RIP olika genomsnittliga livscykelprofiler för deltagande på aktiemarknaden och villkorlig riskandel. Till följd av effekten av cyklisk skevhet på risken med humankapital så innebär HIP-processen mycket mindre heterogenitet i deltagandegraderna mellan människor med olika genomsnittliga inkomstgrader och ett s k fjärilsmönster (butterfly pattern) för den villkorliga riskandelen. Det senare innebär att aktörer med högre genomsnittliga tillväxttakter har en lägre villkorlig riskandel vid tidig ålder jämfört med individer med lågtillväxttakt och de kommer ikapp vid ungefär fyrtio års ålder när ordningen på detta mönster är omvänd.

När jag jämför de modellgenererade profilerna och deras empiriska motsvarigheter genom att använda svenska administrativa data finner jag att det senare ger aningen starkare stöd för RIP-hypotesen.

I det tredje kapitlet *"Heterogenitet i preferenser och portföljval över förmögenhetsfördelningen" (Preference heterogeneity and portfolio choices over the wealth distribution)*, tillsammans med Markus Kondziella och Zoltán Rácz, visar vi att endogen inkomstrisk till följd av preferensheterogenitet mellan individer bidrar till att förklara inkomstojämlikhet.

Utgångspunkten i uppsatsen är att två frågor kvarstår som obesvarade inom hushålls- och makroekonomi, nämligen att förklara individers portföljval respektive höga inkomstojämlikhet. Vi visar att en sammanlänkning av de två litteraturerna kan underlätta att båda frågorna hanteras samtidigt.

Specifikt lägger vi till endogent portföljval till den vanliga inkompletta marknadsmodellen, en avkastningsprocess som inte är normalfördelad, cyklisk skevhet i chocker i arbetsinkomst, Epstein-

Zin preferenser och preferensheterogenitet inklusive heterogenitet i individers tidspreferensgrad, riskaversion och förmåga till eller benägenhet för portföljdiversifiering.

När vi beräknar de modellparametrar som styr heterogeniteten i preferenser för att matcha den ökade riskandelen, deltagandegraden och andelen idiosynkratisk avkastningsrisk över förmögenhetsfördelningen dokumenterad i data, finner vi att för att få en bra förklaringskraft så måste samhället befolkas av två typer av aktörer: en som uppvisar en högre riskaversion och otålighet och den andra karaktäriserad av lägre värden på samma parametrar. I själva verket, då aktörer av den senare typen endogent slutligen hamnar högst upp i fördelningen, kan modellen fånga den positiva korrelationen mellan förmögenhet och riskandel som man funnit empiriskt och i sin tur skapar inkomstojämlikhet i toppen.

När vi studerar resultaten i vår riktmärkesspecifikation med kontrafaktiska samhällen där vi stänger ner olika komponenter en i taget finner vi att modellen alltid förlorar betydligt i förklaringskraft.

IIES MONOGRAPH SERIES

1. Michaely, Michael, *The Theory of Commercial Policy: Trade and Protection*, 1973
2. Söderström, Hans Tson, *Studies in the Microdynamics of Production and Productivity Change*, 1974
3. Hamilton, Carl B., *Project Analysis in the Rural Sector with Special Reference to the Evaluation of Labour Cost*, 1974
4. Nyberg, Lars and Staffan Viotti, *A Control Systems Approach to Macroeconomic Theory and Policy in the Open Economy*, 1975
5. Myhrman, Johan, *Monetary Policy in Open Economies*, 1975
6. Krauss, Melvyn, *International Trade and Economic Welfare,* 1975
7. Wihlborg, Clas, *Capital Market Integration and Monetary Policy under Different Exchange Rate Regimes,* 1976
8. Svensson, Lars E.O., *On Competitive Markets and Intertemporal Resources Allocation*, 1976
9. Yeats, Alexander J., *Trade Barriers Facing Developing Countries*, 1978
10. Calmfors, Lars, *Prices, Wages and Employment in the Open Economy*, 1978
11. Kornai, Jànos, *Economics of Shortage, Vols I and II*, 1979
12. Flam, Harry, *Growth, Allocation and Trade in Sweden. An Empirical Application of the Heckscher-Ohlin Theory*, 1981
13. Persson, Torsten, *Studies of Alternative Exchange Rate Systems. An Intertemporal General Equilibrium Approach*, 1982
14. Erzan, Refik, *Turkey's Comparative Advantage, Production and Trade Patterns in Manufactures. An Application of the Factor Proportions Hypothesis with Some Qualifications*, 1983
15. Horn af Rantzien, Henrik, *Imperfect Competition in Models of Wage Formation and International Trade*, 1983
16. Nandakumar, Parameswar, *Macroeconomic Effects of Supply Side Policies and Disturbances in Open Economies*, 1985
17. Sellin, Peter, *Asset Pricing and Portfolio Choice with International Investment Barriers*, 1990

18. Werner, Ingrid, *International Capital Markets: Controls, Taxes and Resources Allocation*, 1990
19. Svedberg, Peter, *Poverty and Undernutrition in Sub-Saharan Africa: Theory, Evidence, Policy*, 1991
20. Nordström, Håkan, *Studies in Trade Policy and Economic Growth*, 1992
21. Hassler, John, Petter Lundvik, Torsten Persson, and Paul Söderlind, *The Swedish Business Cycle: Stylized facts over 130 years*, 1992
22. Lundvik, Petter, *Business Cycles and Growth*, 1992
23. Söderlind, Paul, *Essays in Exchange Rates, Business Cycles and Growth*, 1993
24. Hassler, John A.A., *Effects of Variations in Risk on Demand and Measures of Business Cycle Comovements*, 1994
25. Daltung, Sonja, *Risk, Efficiency, and Regulation of Banks*, 1994
26. Lindberg, Hans, *Exchange Rates: Target Zones, Interventions and Regime Collapses*, 1994
27. Stennek, Johan, *Essays on Information-Processing and Competition*, 1994
28. Jonsson, Gunnar, *Institutions and Incentives in Monetary and Fiscal Policy*, 1995
29. Dahlquist, Magnus, *Essays on the Term Structure of Interest Rates and Monetary Policy*, 1995
30. Svensson, Jakob, *Political Economy and Macroeconomics: On Foreign Aid and Development*, 1996
31. Blix, Mårten, *Rational Expectations and Regime Shifts in Macroeconometrics*, 1997
32. Lagerlöf, Nils-Petter, *Intergenerational Transfers and Altruism*, 1997
33. Klein, Paul, *Papers on the Macroeconomics of Fiscal Policy*, 1997
34. Jonsson, Magnus, *Studies in Business Cycles*, 1997
35. Persson, Lars, *Asset Ownership in Imperfectly Competitive Markets*, 1998
36. Persson, Joakim, *Essays on Economic Growth*, 1998

37. Domeij, David, *Essays on Optimal Taxation and Indeterminacy*, 1998
38. Flodén, Martin, *Essays on Dynamic Macroeconomics*, 1999
39. Tangerås, Thomas, *Essays in Economics and Politics: Regulation, Elections and International Conflict*, 2000
40. Lidbom, Per Pettersson, *Elections, Party Politics and Economic Policy*, 2000
41. Vestin, David, *Essays on Monetary Policy*, 2001
42. Olofsgård, Anders, *Essays on Interregional and International Political Economics*, 2001
43. Johansson, Åsa, *Essays on Macroeconomic Fluctuations and Nominal Wage Rigidity*, 2002
44. Groth, Charlotta, *Topics on Monetary Policy*, 2002
45. Gancia, Gino A., *Essays on Growth, Trade and Inequality*, 2003
46. Harstad, Bård, *Organizing Cooperation: Bargaining, Voting and Control*, 2003
47. Kohlscheen, Emanuel, *Essays on Debts and Constitutions*, 2004
48. Olovsson, Conny, *Essays on Dynamic Macroeconomics*, 2004
49. Stavlöt, Ulrika, *Essays on Culture and Trade*, 2005
50. Herzing, Mathias, *Essays on Uncertainty and Escape in Trade Agreements*, 2005
51. Bonfiglioli, Alessandra, *Essays on Financial Markets and Macroeconomics*, 2005
52. Pienaar, Natalie, *Economic Applications of Product Quality Regulation in WTO Trade Agreements*, 2005
53. Song, Zheng, *Essays on Dynamic Political Economy*, 2005
54. Eisensee, Thomas, *Essays on Public Finance: Retirement Behavior and Disaster Relief*, 2005
55. Favara, Giovanni, *Credit and Finance in the Macroeconomy*, 2006
56. Björkman, Martina, *Essays on Empirical Development Economics: Education, Health and Gender*, 2006
57. Larsson, Anna, *Real Effects of Monetary Regimes*, 2007

58. Prado, Jr., Jose Mauricio, *Essays on Public Macroeconomic Policy*, 2007
59. Tonin, Mirco, *Essays on Labor Market Structures and Policies*, 2007
60. Queijo von Heideken, Virginia, *Essays on Monetary Policy and Asset Markets*, 2007
61. Finocchiaro, Daria, *Essays on Macroeconomics*, 2007
62. Waisman, Gisela, *Essays on Discrimination and Corruption*, 2008
63. Holte, Martin Bech, *Essays on Incentives and Leadership*, 2008
64. Damsgaard, Erika Färnstrand, *Essays on Technological Choice and Spillovers*, 2008
65. Fredriksson, Anders, *Bureaucracy, Informality and Taxation: Essays in Development Economics and Public Finance*, 2009
66. Folke, Olle, *Parties, Power and Patronage*, 2010
67. Drott, David Yanagizawa, *Information, Markets and Conflict*, 2010
68. Meyersson, Erik, *Religion, Politics, and Development*, 2010
69. Klingelhofer, Jan, *Models of Electoral Competition*, 2010
70. Perrotta, Maria Carmela, *Aid, Education and Development*, 2010
71. Caldara, Dario, *Essays on Empirical Macroeconomics*, 2011
72. Mueller, Andreas, *Business Cycles, Unemployment and Job Search*, 2011
73. von Below, David, *Essays in Climate and Labour Economics*, 2011
74. Gars, Johan, *Essays on the Microeconomics of Climate Change*, 2012
75. Spiro, Daniel, *Some Aspects of Resource and Behavioral Economics*, 2012
76. Ge, Jinfeng, *Essays on Macroeconomics and Political Economy*, 2012
77. Li, Yinan, *Institutions, Political Cycles and Corruption*, 2012
78. Håkanson, Christina, *Changes in Workplaces and Careers*, 2013
79. Qin, Bei, *Essays on Empirical Development and Political Economics*, 2013

80. Jia, Ruixue, *Essays on the Political Economics of China's Development*, 2013
81. Campa, Pamela, *Media Influence on Pollution and Gender Equality*, 2013
82. Seim, David, *Essays on Public, Political and Labor Economics*, 2013
83. Shifa, Abdulaziz B., *Essays on Growth, Political Economy and Development*, 2013
84. Panetti, Ettore, *Essays on the Economics of Banks and Markets*, 2013
85. Schmitt, Alex, *Beyond Pigou: Climate Change Mitigation, Policy Making and Decisions*, 2014
86. Rogall, Thorsten, *The Economics of Genocide and War*, 2015
87. Baltrunaite, Audinga, *Political Economics of Special Interests and Gender*, 2016
88. Harbo Hansen, Niels-Jakob, *Jobs, Unemployment, and Macroeconomic Transmission*, 2016
89. Stryjan, Miri, *Essays on Development Policy and the Political Economy of Conflict*, 2016
90. Karadja, Mounir, *On the Economics and Politics of Mobility*, 2016
91. Kitamura, Shuhei, *Land, Power and Technology, Essays on Political Economy and Historical Development*, 2016
92. Malmberg, Hannes, *Human Capital in Development Accounting and Other Essays in Economics*, 2017
93. Öberg, Erik, *On Money and Consumption*, 2017
94. Lane, Nathaniel, *States of Development: Essays on the Political Economy of Development in Asia*, 2017
95. Prawitz, Erik, *On the Move: Essays on the Economic and Political Development of Sweden*, 2017
96. Iwanowsky, Mathias, *Essays in Development and Political Economics*, 2018
97. Dehdari, Sirus Håfström, *Radical Right, Identity, and Retaliation*, 2018
98. Harmenberg, Karl, *Essays on Income Risk and Inequality*, 2018

99. Kilström, Matilda, *Households' Responses to Policy in Labor and Credit Markets*, 2018
100. Meriläinen, Jaakko, *Essays in Political Economics*, 2019
101. Sigurdsson, Jósef , *Essays on Labor Supply and Adjustment Frictions*, 2019
102. Cocciolo, Serena, *Participatory Governance and Public Service Provision* , 2019
103. Mitrunen, Matti, *Essays on the Political Economy of Development*, 2019
104. Olsson, Jonna, *Work, wealth, and well-being: Essays in macroeconomics*, 2019
105. Darougheh, Saman, *Search and Mismatch*, 2019
106. van Vlokhoven, Has, *On the Cost of Capital, Profits and the Diffusion of Ideas*, 2020
107. Foltyn, Richard, *Essays in Macroeconomics and Household Finance*, 2020
108. Lerva, Benedetta, *Worms, Farms and Schools: Three Essays on Investment Choices in Uganda*, 2020
109. Kragh-Sørensen, Kasper, *On Housing, Mortgages, and Taxation*, 2020
110. Karlman, Markus, *Essays on Housing: tax treatment, prices, and macroeconomic implications*, 2020
111. Kinnerud, Karin, *Financial Choice and Public Policy*, 2020
112. Ghisolfi, Selene, *Fairness, technology adoption, water sanitation and pandemic control: Six essays on four topics in Development Economics*, 2020
113. Åhl, Magnus, *Essays on Individual-Level Wage Stickiness and Forward Guidance*, 2020
114. Dev, Divya *Essays on Gender, Development and Political Economy*, 2021
115. Gallegos Dago, José Elias *Essays on Macroeconomics*, 2022
116. Sun, Xueping *Essays on China's Economic Development: Innovation, Public Debt and Social Connections*, 2022

117. Kramer, John Vincent *Essays on Macroeconomics, Monetary Policy and Mobility*, 2022

118. Almgren, Mattias *Essays on Home Production, Mobility, and Monetary Policy*, 2023

119. Thürwächter, Claire *Essays on Macroeconomics, Monetary Policy and Firm Heterogeneity*, 2023

120. von Carnap, Tillmann *Markets and marketplaces: Essays on access and transformation in remote rural economies*, 2023

121. Peters, Markus *Essays on Savings Behavior, Inflation Measurement, and Growth*, 2023

122. Loiacono, Francesco *Firms and Labor Markets: Essays in Development Economics*, 2023

123. Tebbe, Sebastian *Externalities and Coordination Failures*, 2023

This thesis consists of three self-contained essays focused on how agents balance their exposure to income risk.

Business cycle asymmetry of earnings pass-through analyzes how endogenous income risk emerges from the optimal risk-sharing allocation between workers and firms.

Inferring income properties from portfolio choices shows that endogenous income risk coming from agents' portfolio choices reveals their true labor income process.

Preference heterogeneity and portfolio choices over the wealth distribution illustrates that endogenous income risk ensuing from preference heterogeneity across individuals helps explain wealth inequality.

Gualtiero Azzalini
holds a B.Sc. and a M.Sc. in Economic and Social Sciences from Bocconi University. His research focuses on income risk and its impact on agents' decisions and the macroeconomy.

ISBN 978-91-8014-368-4
ISSN 0346-6892

Department of Economics

Stockholm University