Essays on Macroeconomics and Inequality

Inauguraldissertation
zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaften
der Universität Mannheim

im Herbst-/Wintersemester 2021 vorgelegt von

Johannes Wacks

Abteilungssprecher:	Prof. Volker Nocke, Ph.D.
Vorsitzender der Disputation:	Prof. Klaus Adam, Ph.D.
Referent:	Prof. Tom Krebs, Ph.D.
Korreferent:	Prof. Dr. Matthias Meier
Tag der Disputation:	13. Dezember 2021

Eidesstattliche Erklärung

Hiermit erkläre ich, dass ich die Arbeit selbstständig angefertigt und die benutzten Hilfsmittel vollständig und deutlich angegeben habe.

Frankfurt am Main, am 18. Dezember 2021

Johannes Wacks

Curriculum Vitae

2016–2021	University of Mannheim (Germany)
	Ph.D. in Economics
2014–2016	University of Mannheim (Germany)
	Master of Science in Economics
2012	University of Sevilla (Spain)
	Visiting Student
2010–2014	University of Göttingen (Germany)
	Bachelor of Arts in Economics

Acknowledgments

I am deeply indebted to my advisor, Tom Krebs, for invaluable guidance and support over the recent years. His prior research inspired my interest in understanding the causes and consequences of income and wealth inequality. Discussions with him greatly improved my own research. I am grateful to Matthias Meier, who encouraged me to take the first steps in my own research projects at very early stages of the PhD program, and who was a continuing source of valuable feedback throughout the past years.

I thank Klaus Adam for the many times he gave me thoughtful feedback on my work. His doctoral colloquium was the ideal environment to exchange views on first research ideas. My work has also greatly benefited from discussions with Antoine Camous, Husnu Dalgic, Andreas Gulyas, Anne Hannusch, David Koll, Krzysztof Pytka, Michèle Tertilt, Minchul Yum and participants at the Mannheim Macro Seminar. Benjamin Moll's fascinating research, his presentation in the Department Seminar in 2017, and the EABCN Training School he taught in 2018, provided inspiration for my own research. I have further received valuable comments on my work from Pedro Amaral, Mark Bils, Elizabeth Caucutt, Juan Carlos Conesa, Matias Cortes, Tim Kehoe, Kurt Mitman, José-Víctor Ríos-Rull, and Ayşegül Şahin.

I especially want to thank my amazing co-author, Alexander Matusche. From the first day of the PhD program, where we started out as office mates, I enjoyed every second of being colleagues and friends. I cannot count the number of insightful discussions we had and the times that he helped me structure my thoughts more clearly. I consider myself lucky to have been part of such a great cohort of PhD students. Alex, Johannes, Timo, Karl, Frederick, Michaela, Xi, Mark, Sebastian, Max, thanks for the fun times we had at the office, your ideas on how to improve my research, and for making the five years in graduate school an unforgettable time. I also thank Florian Jaster, Marco Matschewski, Daniel Runge, Ole Schmidt and Jan Sun for valuable feedback on my dissertation.

I am very appreciative of the administrative staff at the University of Mannheim, in particular Corinna Jann-Grahovac, Birgit Junker, Sylvia Rosenkranz, Susanne Putz, Marion Lehnert, Caroline Mohr, Ulrich Kehl, and Dagmar Röttsches. They have enabled me to focus on my research, and made being at the office very enjoyable. I am grateful to the University of Mannheim, the Graduate School of Economics and Social Sciences (GESS) and the foundation Geld und Währung for financial support, and the state of Baden-Württemberg for computational resources through bwHPC.

Lastly, I want to thank my family. My parents, grandparents and sisters have supported me and cheered me up through all these years and I couldn't be more grateful to have all of them in my life. Most importantly, I thank my wife Nora. If I was ever overwhelmed by the challenges of the PhD program, she was there for me and made me smile. At every step of the journey she stood by my side. Thank you for your optimism, for your happiness, for your love.

Contents

Preface

This dissertation consists of three self-contained chapters. The common thread that connects these chapters is the question how household heterogeneity, in particular inequality in income and wealth, affects macroeconomic dynamics.

Chapter 1: Labor Market Polarization with Hand-to-Mouth Households I argue that borrowing constraints are important for understanding the welfare and output consequences of labor market polarization. Using data from the PSID and SCF, I document that i) a large fraction of routine workers who switched to manual and abstract occupations suffered short-term wage losses, ii) one third of routine workers are hand-to-mouth (hold few liquid assets), and iii) being hand-to-mouth predicts a low probability of leaving routine occupations. I build a general equilibrium incomplete markets model featuring three occupations, a continuum of skill types, occupation-specific human capital and a realistic share of hand-to-mouth households. A fall in the price of capital causes the routine wage to decline relative to wages in the other occupations. The presence of a large share of households who are close to the borrowing constraint and thus reluctant to invest in future wage growth significantly protracts the reallocation of labor away from the declining routine occupation. I study two policies that alleviate the borrowing constraint upon switching the occupation. Both a government loan program and the payment of transfers raise social welfare and output, and speed up the reallocation of workers away from the declining routine occupation. While disadvantageous for the high-skilled, the policies benefit medium- and low-skilled workers.

Chapter 2: Does Wealth Inequality Affect the Transmission of Monetary Policy? This chapter is joint work with Alexander Matusche.[1] We provide empirical evidence that higher wealth inequality between households is associated with stronger real effects of monetary policy. We use three separate data sets and different empirical strategies. First, we employ state-dependent local projections to show that the US and the UK exhibited stronger real effects of monetary policy in times of higher wealth inequality. Second, we construct a measure of wealth inequality at the US state level from estate tax data and use it to document that output and unemployment respond more strongly to interest rate changes in states with higher wealth inequality. Third, we use HFCS micro data to estimate wealth

[1] University of Mannheim.

inequality in Euro Area countries. We find that ECB monetary policy has stronger effects on economic activity in countries where wealth is distributed more unequally. All three pieces of evidence thus point to the conclusion that higher wealth inequality is associated with stronger effects of monetary policy.

Chapter 3: Monetary Policy and Wealth Inequality—The Role of Entrepreneurs This chapter is joint work with Alexander Matusche. We show that rising inequality in US household wealth holdings has come with a shift of wealth from workers to entrepreneurs and ask how this affects the transmission of monetary policy to the real economy, in particular to aggregate investment. We develop a Heterogeneous Agent New Keynesian model in which some households are entrepreneurs who can invest in their private firm with risky returns. In response to expansionary monetary policy, entrepreneurs rebalance their portfolios and expand investment into their firm. The model matches the distribution of returns from private businesses over owners' net worth observed in the Survey of Consumer Finances. This is important because a lower excess return over the risk-free rate leads to stronger portfolio rebalancing towards the private business. Our model attributes a quantitatively important role to entrepreneurs in the transmission of monetary policy. If entrepreneurs do not react to the change in the interest rate, the output response is about 50% smaller. An increase in wealth inequality, modeled as a shift of wealth from workers to entrepreneurs, strengthens the effects of monetary policy. An increase in the top 10% wealth share by one percentage point implies that the aggregate output response to a rate cut increases by 3% to 20%.

Chapter 1

Labor Market Polarization with Hand-to-Mouth Households

1 Introduction

Technological change has had a major impact on labor markets in recent decades. Many jobs that were previously performed by humans have become automated and are now completed by machines. This trend has benefited workers in some occupations while harming those in others. On the one hand, jobs intensive in routine, i.e. easily codifiable tasks (e.g. assembly line workers, bookkeepers), have increasingly become automated. On the other hand, jobs intensive in manual tasks (e.g. taking care of the elderly or children) and abstract tasks (e.g. teachers, managers) have remained and often profited from technological change, as they are difficult to master for machines. Consequently, wages and employment shares in occupations intensive in routine tasks have declined in recent decades compared to those in manual and abstract occupations.[1]

In this paper, I argue that borrowing constraints and the ability to smooth consumption are important to understand the output and welfare consequences of this technological change and the labor market reallocation it has caused. To share in the benefits of technological growth, many routine workers have left their old jobs and switched into occupations in which wages have been rising (Cortes, 2016). I show that these switches, while providing benefits in terms of higher wages in the medium to long term, were often accompanied by initial wage losses. This pattern, which can be rationalized by the presence of occupation-specific human capital, makes occupational choice a dynamic investment decision. The ability to smooth consumption, i.e. the distance from the borrowing constraint, therefore becomes an important determinant of whether a worker decides to switch or stay in her old job.

The main contribution of this paper is to embed routine-biased technological growth into a general equilibrium model of the US economy featuring incomplete markets and occupation-

[1] These empirical patterns have been documented, among others, by Acemoglu and Autor (2011), Autor and Dorn (2013), Autor, Levy, et al. (2003), Cortes, Jaimovich, and Siu (2017), and Goos et al. (2014).

specific human capital. I use the model to demonstrate that borrowing constraints play an important role in shaping the reallocation of labor away from jobs in routine occupations. The aggregate labor market transition is impeded by the presence of hand-to-mouth households, i.e. those with few or no liquid assets, as they are unwilling to invest in future earnings growth at the cost of bearing short-term wage losses. This gives rise to an inefficiency, as the discounted gains from switching can outweigh the initial costs. Building on this insight, I use the model to ask how government policies aimed at alleviating the borrowing constraint of occupational switchers affect output and welfare. In particular, I study the introduction of a government loan program and the payment of transfers, both targeted at formerly routine workers who have left their occupation.

To assess whether the friction I propose is empirically relevant, I build on and extend previous empirical findings. Using panel data from the Panel Study of Income Dynamics (PSID), Cortes (2016) shows that, compared to those workers who stayed in routine occupations, workers who switched to either abstract or manual occupations saw faster wage growth and thus enjoyed long-term wage gains. Adding to this, I show that a large share of switchers did, however, suffer wage losses in the short run, i.e. in the year immediately following the switch. Next, I use data from the Survey of Consumer Finances (SCF) to demonstrate that about a third of routine workers fall under common definitions of being hand-to-mouth, i.e. hold very few liquid assets. This shows that potential short-term wage losses might have posed an important obstacle to an occupational switch for a large fraction of routine workers. Lastly, I bring these two pieces of information together, turning again to recent waves of the PSID, which provide information on asset holdings. I ask whether the current hand-to-mouth status is predictive of leaving the routine occupations for a new job, and find that, in line with the mechanism described above, hand-to-mouth households are less likely to make such a switch.

I then build a heterogeneous agent model with idiosyncratic labor income risk and incomplete markets, that takes account of these empirical findings. On the supply side, there exists a representative firm that uses two types of capital (Information and Communication Technology (ICT) and non-ICT capital) and three types of labor (manual, routine and abstract labor) as inputs. The driving force of technological change is an exogenous fall in the price of ICT capital relative to the consumption good. The prices of all production factors are endogenously determined within the model. On the demand side the model features households that die stochastically and are then replaced by newborns (perpetual youth). Households are heterogeneous in several ways. They have a fixed skill type, which affects their optimal occupational choice. In line with the data, low-skill types work in low-wage manual jobs, medium-skilled in routine jobs in the middle of the wage distribution, and high-skilled workers sort into the high-paying abstract occupation. While working in either of the three occupations, households accumulate occupation-specific human capital, which depreciates once they switch to another occupation. On top of this, households are exposed to uninsurable idiosyncratic productivity risk, causing them to engage in precautionary saving. Households can use two assets for saving, a liquid and an illiquid one. This enables the model to generate a share of hand-to-mouth agents that matches the data.

I calibrate the model to the US economy, targeting the employment shares in each occupation and the share of income accruing to labor in 1980 and 2020, the share of hand-to-mouth households, as well as the average wage changes of routine workers who switch to either the abstract or to the manual occupation. I solve for the steady state when the price of ICT capital is relatively high (representing 1980) and when it is relatively low (representing today), and compute the perfect foresight transition path between the two steady states. The calibrated production function implies that ICT capital is relatively easy to substitute with routine work and relatively complementary to manual and especially abstract work. This leads to a lower routine employment share in the new steady state compared to the old one.

Consistent with the empirical evidence and the proposed mechanism the model predicts an important role of liquid assets and hence the ability to smooth consumption for the switching behavior of routine workers. I demonstrate this, first, by zooming in on individual policy functions. Conditional on skill and occupation-specific human capital, routine workers leave the occupation earlier if they possess liquid asset savings than if they are hand-to-mouth. Second, I conduct counterfactual simulations of the economy in which I assume that all households choose their occupation like the ones who are well insured against shocks, in particular those with liquid assets above the 70th (90th) percentile of the liquid asset distribution. In these counterfactuals, the employment share in the abstract occupation would have been two (three) percentage points higher on average along the transition than in the baseline. Third, I conduct the same regressions of hand-to-mouth status on future switching decisions as before in the PSID, only now in a synthetic panel of households simulated from the model, and I find the same negative association as in the data.

Next I ask how the policymaker can address the inefficiency introduced by the borrowing constraint by studying policies that target routine workers who leave their former occupation. First, I study the introduction of a government loan program, which provides switchers with additional liquidity upon leaving the routine occupation. Second, I let the government pay a transfer to the switchers, partly covering the temporary wage loss they endure while building up occupation-specific human capital in their new occupation. Both policies are financed via distortionary labor income taxation. Focusing on policies targeting only the routine to abstract switchers, I find that both programs increase aggregate welfare. They do so both because households who become eligible directly benefit from the programs, and because of general equilibrium effects. As more workers switch into the abstract occupation when the policies are in place, wages of (rich) abstract workers decline and wages of (poor) manual and routine workers rise. Quantitatively, medium-skilled types can increase expected lifetime consumption by 0.3% (1.0%) under the optimal loan (transfer) program. Besides these effects on welfare, the policies also cause output to rise. This is driven by the crowding-in of ICT capital, which increases by up to 1.0% (3.0%) compared to the baseline transition, and by an increase in average labor productivity.

The rest of the paper is structured as follows. I first discuss the related literature. In Section 2 I use a small, highly stylized model of dynamic occupational choice to develop the intuition for the key mechanism. In Section 3 I present empirical evidence that motivates the quantitative model. The full quantitative model is presented in Section 4. In Section 5 I study the transition between two steady states, while in Section 6 I conduct the policy experiments. Section 7 concludes.

Related Literature A large empirical literature has explored the phenomenon of labor market polarization. Acemoglu and Autor (2011) point out that a notable polarization both of wage growth and of jobs has taken place in the US since the 1980s. Plotted across hourly wage quantile, wage changes exhibit a u-shape, featuring higher growth at low and high quantiles. Similarly, both the share of employment in high skill, high wage occupations and low skill, low wage occupations grew over this time frame. In a theoretical model they link these patterns to the "routinization" hypothesis, i.e. that routine jobs of the income middle class have been increasingly automated, as in Autor, Levy, et al. (2003). Building on these insights, Autor and Dorn (2013) point out that the relative rise in earnings and employment at the bottom of the occupational skill distribution in the US can be attributed to the rise of service occupations. I add to this literature by showing that the presence of borrowing constraints and imperfect consumption smoothing have slowed down the reallocation of labor out of routine occupations.

Cortes, Jaimovich, Nekarda, et al. (2020) find that a drop in the inflow rates from non- and unemployment into routine occupations explains 34–43% of the decline in routine labor over recent decades, with young cohorts playing a disproportionate role.[2] They also find that outflows to non- and unemployment are not important for explaining the decline. While I abstract from non- and unemployment in my model, its perpetual youth structure allows reduced entry of new labor market entrants into routine occupations to play a role that is quantitatively in line with Cortes, Jaimovich, Nekarda, et al. (2020) in bringing down the aggregate routine employment share. That said, the authors note that 54–60% of the drop in routine employment cannot be explained by changed exit and entry to non- and unemployment, and is therefore due to job-to-job transitions and other transitions not covered by their decomposition. In this paper, I focus on job-to-job transitions and reduced inflows from new labor market entrants as the sole drivers of the decline in routine employment.

Cortes (2016) proposes a simple theoretical model, in which workers differ by fixed skill type. Low types sort into manual, medium into routine and high-skilled types into abstract occupations. He finds support for this assumption of a "catch-all" skill and the proposed ordering of the occupational groups using data from the PSID. He shows that low-(high-) skilled routine workers were most likely to exit to the manual (abstract) occupations while

[2] This is also in line with the findings in Adão et al. (2020). They show that employment responses to technological innovations are to a large extent driven by new generations making different skill investment choices than older cohorts before entering the labor market.

the labor market polarized.[3] Importantly, he finds that compared to routine workers who stayed in their occupation, those who switched to either the abstract or manual occupations saw faster wage growth. I extend these empirical results by documenting that many routine switchers faced short-term wage losses upon changing their occupation and that liquid asset holdings were a strong predictor of the switching decision.

There has also been recent progress in embedding labor market polarization into quantitative macroeconomic models. My model is closest to vom Lehn (2020). Households in his model are heterogeneous in skill type and endogenously sort into the three broad occupational groups as in Cortes (2016). While in vom Lehn (2020) the household sector can be represented by a representative agent, in my model workers face uninsurable idiosyncratic income risk (Aiyagari, 1994; Bewley, 1983; Huggett, 1993). I also add occupation-specific human capital and an intensive labor supply margin. The latter ensures that progressive labor income taxation has a dampening effect on the aggregate labor supply. This is important when studying the effects of labor market policies that are financed by raising labor taxes, as in Section 6.[4]

Moll et al. (2021) add household heterogeneity in skills and dissipation shocks to wealth accumulation to the task-based framework developed in Acemoglu and Restrepo (2018b). They simulate a trend in automation to study its implications for wealth inequality, and find that automation, by driving up the interest rate on capital, leads to higher wealth inequality. I differ from Moll et al. (2021) in building a model with uninsurable idiosyncratic income risk and occupation-specific human capital, and by focusing on the normative aspects of labor market polarization. The interest in policy analysis is shared by Jaimovich, Saporta-Eksten, et al. (2021). However, while they put the focus on labor market frictions, I zoom in on the importance of borrowing constraints along the transition path. As a result, the policies studied in Section 6 are tailored towards alleviating the borrowing constraint of routine workers who decide to leave their old occupation.

That human capital is at least in part tied to a worker's specific occupation, or the tasks performed in them, has been documented in several empirical studies (Cortes and Gallipoli, 2018; Gathmann and Schönberg, 2010; Kambourov and Manovskii, 2009b; Sullivan, 2010). Two recent quantitative models of polarization, Dvorkin and Monge-Naranjo (2019) and Kikuchi and Kitao (2020), capture occupation-specific capital. While Dvorkin and Monge-Naranjo (2019) do not model a consumption-saving choice, precluding the mechanism proposed here, Kikuchi and Kitao (2020) model such a choice and obtain an endogenous wealth distribution as I do. The focus of their study, however, lies on the welfare effects of labor market polarization on different age and skill groups, and not on analyzing the interaction between borrowing constraints and occupational choices. Also, they assume wages to be exogenous, while all factor prices are endogenous in my model.

[3] Bachmann et al. (2019), Bessen et al. (2020), Dauth et al. (2021), and Edin et al. (2021), among others, provide empirical evidence on how workers in labor markets other than the US were affected by automation.

[4] I further deviate from vom Lehn (2020) in that I assume that households differ ex ante only in general skill type. In vom Lehn (2020), households draw three additional skill types, one for each occupation.

2 A Small Model of Dynamic Occupational Choice

In this section I illustrate the core mechanism at work in the full model of Section 4. I highlight that in a dynamic setting, in which relative wages between occupations change over time, a worker potentially makes differing occupational choices depending on whether she is at the borrowing constraint or not. A necessary ingredient to the model for this channel to be active is that human capital is occupation-specific, as in this case occupational choice becomes an investment decision. A trade-off arises between building up human capital in the occupation where wages are growing and the desire to smooth consumption.

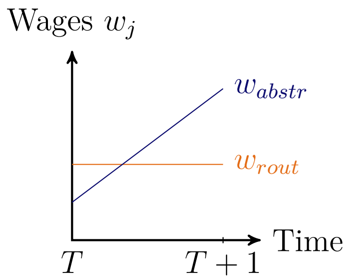


Figure 1.1: Wage paths

The model has two time periods, T and $T+1$. There are two occupations, routine and abstract. The wage per efficiency unit of labor in the routine occupation is one in both periods, i.e. $w_{r,T} = w_{r,T+1} = 1$. The abstract wages are $w_{a,T} = \frac{1}{\omega}$ and $w_{a,T+1} = \omega$, with $\omega > 1$ (Figure 1.1).[5] This captures the fact that while in the 1980s it paid more for certain workers to have a job at an assembly line or to be a bookkeeper than to be a manager or a teacher, in the 1990s this ordering had reversed due to an increase in abstract wages.

Consider a household who makes an occupational choice at the beginning of period T and again at the beginning of period $T+1$. Her income in every period is $y_t = w_{j,t} \cdot h_t$, where $h_t \in \{\underline{h}, \bar{h}\}$ is occupation-specific human capital, with $0 < \underline{h} < \bar{h}$. If the household works in the same occupation j in T and $T+1$, she is an experienced worker in occupation j in $T+1$ with certainty ($h_{T+1} = \bar{h}$). An occupational switch at the beginning of period t leads to full depreciation of human capital ($h_t = \underline{h}$). The household values consumption, with $u(c) = \ln(c)$. I assume that the household holds zero assets and that she discounts future utility by a factor $\beta \in (0, 1)$.

Suppose that at the beginning of period T the household has not yet gathered experience in either of the two occupations. What is her optimal occupational choice? Assuming first that the household can freely borrow against future income, her decision problem is:

[5] This symmetry is not crucial for the results but simplifies the algebra. Appendix B relaxes this assumption.

$$\max_{c_T, c_{T+1}, j_T, j_{T+1}} \ln c_T + \beta \ln c_{T+1}$$
$$\text{s.t.: } c_T + \frac{c_{T+1}}{1+r} = y(j_T) + \frac{y(j_T, j_{T+1})}{1+r}$$

where r is the exogenous interest rate, c consumption, and j_t is the occupational choice at the beginning of period t. If instead the household is exogenously prevented from borrowing, in addition it has to hold that

$$c_T \leq y_T \ .$$

Given that she can choose between the two occupations at the beginning of each period there exist four possible combinations of occupational choices, which in turn determine her labor income:

1. $\{j_T = r, j_{T+1} = r\} \rightarrow \{y_T = \underline{h}, y_{T+1} = \bar{h}\}$
2. $\{j_T = r, j_{T+1} = a\} \rightarrow \{y_T = \underline{h}, y_{T+1} = \omega \cdot \underline{h}\}$
3. $\{j_T = a, j_{T+1} = a\} \rightarrow \{y_T = \underline{h}/\omega, y_{T+1} = \omega \cdot \bar{h}\}$
4. $\{j_T = a, j_{T+1} = r\} \rightarrow \{y_T = \underline{h}/\omega, y_{T+1} = \underline{h}\}$

Note that option 4 can be ruled out as an optimal choice, as income in both periods is higher in option 2.

Proposition 1. *If not borrowing-constrained, the household chooses the abstract occupation in $t = T$ iff*

$$1 + r \leq \omega(\bar{h}/\underline{h} - 1) \min \left\{ \frac{\bar{h}/\underline{h}}{\bar{h}/\underline{h} - 1}, \frac{\omega}{\omega - 1} \right\} .$$

In this case, she is a net borrower in $t = T$.
If borrowing-constrained, the household never chooses the abstract occupation in $t = T$.

The proof is relegated to Appendix A. The proposition shows that there exists a set of parameter combinations under which the household optimally chooses option 3 when she is not borrowing-constrained. This means that she works in the abstract occupation in period T, even though wages are still higher in the routine occupation. As a consequence, she borrows against future income in period T.

The condition in Proposition 1 becomes more likely to hold when the interest rate r is low, such that borrowing against future income is cheap, and when the human capital spread $\bar{h}/\underline{h}$ is large, such that starting to gain experience in the abstract occupation already today is valuable. Lastly, due to the non-linear way in which it affects earnings in the two periods in opposite directions, ω has an ambiguous effect on the occupational choice in T.

The second part of Proposition 1 highlights the main inefficiency. Under no combination of parameters will the household work in the abstract occupation in T if she cannot borrow. The intuition is simple: foregoing high wages in the routine occupation today is too costly, and if wages are very high in the abstract occupation in period $T+1$, she can still switch then. That the household always chooses the routine occupation in T when borrowing constraints bind is clearly inefficient: the first result of the proposition shows that for some parameters the long-term gains from switching, in form of discounted future profits, outweigh the short-term costs. In these cases, choosing the abstract occupation in T is a profitable investment.

In Appendix A I show that a similar proposition holds for a worker who instead of being inexperienced in period T is an experienced routine worker, i.e. $h_T = \bar{h}$. Again, for some parameter combinations and when borrowing is allowed it is optimal to switch to the abstract occupation in T, even though this entails a short-run earnings loss. In case of binding borrowing constraints, however, experienced routine workers stay in the routine occupation in T and make the switch to abstract only in period $T+1$, if at all.

Discussion and extensions Appendix A.2 contains an extension of the simple model with time-varying aggregate productivity. This extended model can qualitatively account for the fact that a higher share of households exits the routine occupations during a recession (because the opportunity costs of doing so are smaller than during a boom). This is in line with empirical evidence in Hershbein and Kahn (2018) and Jaimovich and Siu (2020).

Appendix B extends the simple model in two ways. It endogenizes wages and introduces a continuum of households who are heterogeneous with respect to their skill type. This extended set-up allows me to characterize analytically the socially optimal allocation of labor across the two occupations when the abstract wage grows over time relative to the routine wage. I can show that, in line with the intuition conveyed above, in an economy in which households are hand-to-mouth, too little labor is supplied in the abstract occupation compared to the first-best. Hence, a policy that alleviates the borrowing constraint might raise output by improving the labor allocation.

The mechanism laid out in the simple model does not depend on, but is similar to the existence of occupation-specific returns to tenure. For instance, returns to tenure might be higher in the abstract than in the routine occupation, in which case the ratio $\bar{h}^j/\underline{h}$ would be occupation-specific, with $\bar{h}^a/\underline{h} > \bar{h}^r/\underline{h}$ (e.g. a manager's occupational experience being more highly rewarded than that of a bookkeeper). Then, even if wages per efficiency unit in the two occupations stayed constant over time, income paths would resemble those depicted for wages in Figure 1.1, and borrowing-constrained households might end up inefficiently choosing the routine occupation in $t = T$. While I allow for heterogeneous tenure profiles in the full model in Section 4, the focus of this paper is technological change and the time-varying demand—and hence time-varying wages per efficiency unit—it induces across occupations. Consequently, the policies studied in Section 6 are aimed at helping experienced routine workers switch out of their declining occupation.[6]

In the simple model I have assumed initial asset holdings to be zero. In reality, households can save in anticipation of future wage changes and occupational switches. The more sophisticated model of Section 4, which endogenizes the wealth distribution, is therefore necessary to judge whether the friction proposed here has had meaningful implications for aggregate variables during recent decades. The full quantitative model also includes three instead of two occupations, speaking to the literature on labor market polarization. Before describing it, though, I present empirical evidence that supports the relevance of the just described mechanisms.

3 Empirical Evidence

I begin this section by reviewing the concept of labor market polarization and introducing the three broad occupational groups, as defined in earlier studies (Autor and Dorn, 2013; Jaimovich and Siu, 2020). Then, I demonstrate that the wage paths depicted in Figure 1.1, i.e. short-term losses but long-run gains from leaving the routine occupation, have been an empirically relevant phenomenon in recent decades. Hence, for many workers switching out of the routine occupation was indeed an investment decision. Adding to Cortes (2016)'s finding that switchers out of routine occupations saw faster wage growth than stayers, using PSID data I show that 45%–60% of these switchers saw initial wage losses. Next, using the SCF, I document that, historically, a significant fraction of about 35% of routine workers has been hand-to-mouth, i.e. holding only very few liquid assets. When interpreted as a proxy for being at or close to the borrowing constraint, this indicates that many people fell in the category of constrained households in the simple model above. Last, I turn to recent waves of the PSID and exploit its panel dimension to show that not being hand-to-mouth has been predictive of leaving the routine occupation.

3.1 Labor market polarization and broad occupational groups

To describe the phenomenon of labor market polarization, Autor and Dorn (2013) first order 318 detailed occupations according to their skill level, which they proxy by the average hourly wage earned in the occupation in 1980. They show that between 1980 and 2005 occupations at the bottom and at the top of the skill distribution gained both in terms of employment shares, and in terms of wages relative to occupations found in the center of the skill distribution (cf. their Figure 1). Scrutinizing the task content of each single occupation using the US Department of Labor's Dictionary of Occupational Titles, they

[6]For a treatment of how returns to college and its cost shape the optimal design of student loan programs in the presence of borrowing constraints, see Lochner and Monge-Naranjo (2012, 2016). In Section 6 I show that the policies raise welfare significantly more along the transition path, which features time-varying relative wages between occupations, than they do in the steady state, in which occupation-specific returns to tenure are present, but relative wages are constant.

divide occupations into six broad groups. For tractability, the literature then often subsumes these further into three broad occupational groups: manual, routine and abstract (Jaimovich and Siu, 2020; vom Lehn, 2020).[7]

The first group of occupations are intensive in **manual** tasks, i.e. they require eye-hand-foot coordination, adapting to new surroundings often, and (non-trivial) interaction with other humans. These occupations can be found at the bottom of the wage distribution and are typically service occupations. Examples for manual occupations are health and nursing aides, workers in child and elderly care, bus drivers, waiters and waitresses, door-to-door/street sales, or janitors and gardeners.[8]

The second group consists of occupations that require **routine** tasks to be performed, i.e. calculations, record-keeping, repetitive customer service, repetitive assembly, picking or sorting. Many occupations in the middle of the wage distribution belong to this group. Typical examples of routine occupations are bookkeepers, accounting clerks, secretaries, bank tellers, as well as machine operators and assemblers or butchers and meat cutters.[9]

Occupations of the third group require **abstract** tasks, i.e. cognitive thinking, forming/testing hypotheses, persuading, managing or organizing. These kind of jobs are found in the high-wage occupations, and typical examples are school teachers, managers, police and detectives, public service (e.g. city planners), or engineers.

The term "polarization" refers to the fact that the occupational wage and employment distributions have shifted towards their poles since the 1980s, while the middle of these distributions, where many of the routine occupations are found, has "hollowed out". The employment share in the routine occupations fell from 58.5% in 1980 to 46.2% in 2005, while it increased from 31.6% to 40.9% in the abstract, and from 9.9% to 12.9% in the manual occupations (Autor and Dorn, 2013). One commonly cited cause for the decline of the routine occupations is technological change that substitutes for routine labor (Autor, Levy, et al., 2003; Jaimovich, Saporta-Eksten, et al., 2021; vom Lehn, 2020). The fact that capital, most prominently ICT capital, has become much cheaper over the recent decades has led to a replacement of tasks formerly performed by humans with machines. Manual and abstract jobs are not so easily substituted by ICT capital, as they require either non-trivial interaction with humans or tasks such as managing and organizing, all of which machines struggle to excel at. Following the literature, a falling relative price of ICT capital will be the driver of technological change in the model of Section 4.

[7]Following Jaimovich and Siu (2020), the services occupations are the only occupations I categorize as manual. See Appendix C for details. The assignment is complete and mutually exclusive: each of the detailed occupations is considered to be part of exactly one of the three broad groups.

[8]For an extensive treatment of the classification procedure and more examples for occupations also refer to Autor, Levy, et al. (2003).

[9]Some studies make a further distinction between routine manual and routine abstract occupations. I abstract from this further division for simplicity.

3.2 Wage paths and liquid asset holdings

Did workers who switched jobs face the wage paths depicted in Figure 1.1, i.e. initially lower wages in the manual or abstract occupation, but long-run wage gains from switching? In this section I provide evidence that this was indeed the case for a significant fraction of workers who left the routine occupations.

Wage paths of routine workers: switchers vs. stayers

Using data from the PSID between 1976 and 2007, Cortes (2016) shows that compared to routine workers who stayed in their occupation, those who switched to manual or abstract occupations experienced faster wage growth after they had switched. In particular, wages of workers who switched to manual occupations were 11.2% lower after one year, but 11.5% higher after ten years, than wages of those who stayed. Wages of routine workers who switched to abstract occupations were higher on average by 3.4% after one year compared to wages of those who stayed, and 16.3% higher after ten years (see Figure 1.13 in the Appendix). This pattern is consistent with a relative wage increase in manual and abstract relative to routine occupations, as in Figure 1.1.

To take a closer look at the distribution of wage changes upon leaving routine occupations, I use PSID waves 1976 to 2017, following Cortes (2016) in the definition of variables and in sample selection. In particular, I focus on employed male household heads, aged 16–64. The only two differences in terms of the sample in my analysis are the extended time period and that, to be consistent throughout this paper, I only categorize low-skilled services occupations as manual occupations, while Cortes (2016) uses a slightly broader definition for his baseline results. I relegate all further details to Appendix C.2.

I find that of all the workers who switched from routine to manual (abstract) occupations from one year to the next and for whom wages are observed in both years, 58% (42%) saw their wage decline upon switching.[10] Hence, Cortes (2016)'s result that the one-year wage change of switchers to the abstract occupation was on average 3.4% higher than that of the stayers masks a lot of heterogeneity. While some moves from routine to abstract might have been due to career advancement, leading to wage gains, almost half of these switches have come with wage losses and are therefore less likely to represent career progression.[11]

The point here is not to claim that only leaving routine occupations can lead to (temporary) wage losses. Both a significant fraction of workers leaving manual (33%) and abstract occupations (48%) saw their wage decline year-on-year. However, as Cortes (2016) shows, workers who switched out of manual and abstract occupations did not experience faster wage growth than stayers, as opposed to routine switchers. The takeaway here is therefore that

[10] Moreover, 61% (46%) of switchers from routine to manual (abstract) saw a wage change smaller than 1.3%, which was the average wage growth of routine stayers in this time period. Note that for this part of the analysis I only use PSID waves up until 1997, after which the PSID was conducted bi-annually.

[11] Mukoyama et al. (2021) provide empirical evidence that in the U.S. almost none of the net reallocation of labor from routine into abstract and manual occupations was driven by within-firm switches.

it is predominantly for the group of routine workers that switching can be interpreted as an investment decision. In Appendix C.2 I document moreover that also the subgroup of routine switchers who faced initial wage losses exhibited faster wage growth than routine stayers.

Hand-to-mouth shares

I have so far shown that the wage pattern depicted in Figure 1.1 has been relevant for many routine workers. But how many routine workers face binding liquidity constraints? In order to address this question, I use liquid asset holdings to proxy for closeness to the borrowing constraint. Recent work by Kaplan, Violante, et al. (2014) has shown that although only about 10% of US households have zero or close to zero net worth ("poor hand-to-mouth"), another 20% hold only very few liquid assets, while possessing some illiquid assets, e.g. a house. I extend these findings, splitting the data by the three broad occupations.

I use the SCF, instead of the PSID, to shed light on hand-to-mouth shares by occupation both because the PSID has started including detailed information on asset holdings only in 1999 and because the focus of the SCF is acquiring accurate information on wealth. I use twelve waves, from 1989 to 2019. In terms of sample selection and classifying households as hand-to-mouth, I follow Kaplan, Violante, et al. (2014). In particular, I consider all households whose head is aged 22–79, and discard those who report negative labor income or whose only positive income stems from self-employment. I then relate each household's liquid assets to current income, and classify it as hand-to-mouth if liquid asset holdings are either zero (or positive but close to zero), or if liquid assets are close to an imputed borrowing constraint, equaling one times monthly income. Appendix C.3 lays out the details.

I find that, averaged across time, 35% of households whose head was currently working in routine occupations were hand-to-mouth. This compares to a share of 40% among manual, and 19% among abstract workers. Appendix C.3 depicts the time series as well as confidence intervals around the estimates. This demonstrates that for a large fraction of routine workers binding borrowing constraints have potentially been an impediment to leaving their occupation, if such a switch entailed short-run wage losses. This might have caused them to stay in the routine occupations for a relatively long time, reminiscent of the behavior of the constrained household in the simple model of Section 2.

Liquid assets and switching behavior

In a last step, I ask whether being hand-to-mouth has had predictive power over the decision to leave the routine for the manual or abstract occupations, as the model of Section 2 suggests. To this end, I turn back to the PSID data, waves 1999 until 2017, because these provide information on assets (at the household level) and because, unlike the SCF, the PSID has a panel dimension and allows me to follow individuals over time. I estimate the following model:

$$\text{switch}_{i,t,t+2} = \beta \cdot HtM_{i,t} + \gamma \cdot X_{i,t} + u_{i,t} \ . \tag{1.1}$$

Here, $\text{switch}_{i,t,t+2}$ is a dummy variable which is equal to zero if worker i was employed in a routine occupation in year t and $t + 2$ and equal to one if the worker was employed in a routine occupation in year t and a manual or abstract occupation in year $t + 2$.[12] I look at two-year differences as the PSID went to a bi-annual frequency in 1997. $HtM_{i,t}$ is a dummy variable indicating whether the household is hand-to-mouth in year t, and $X_{i,t}$ are control variables. $u_{i,t}$ is an exogenous error term with $\mathbb{E}[u_{i,t}] = 0$. I control for tenure in the broad occupation, a dummy indicating the region, age and its square, unionization status, married status, and a linear time trend. As in the quantitative model of the next section innate ability, or skill, will be an important determinant of switching behavior, I include in some specifications a dummy for whether individuals have received at least some years of college education.[13] For the baseline results, in line with Cortes (2016), I only include males in the sample. I cluster standard errors at the individual level.

Table 1.1 shows the results. The first row of columns 1 and 2 show the estimate of β. I find that hand-to-mouth agents are less likely to leave the routine occupation, compared to agents who hold a buffer of liquid assets. The likelihood of switching is 2.1 to 3.0 percentage points smaller for the former group than for the latter, everything else equal. This is in line with the mechanism laid out above, by which more borrowing-constrained households, for fear of temporary earnings losses, delay their move away from the routine occupations. Also, in line with what would be expected, more years of occupational tenure make switching less likely. Skill, as measured by education, has a positive impact on switching probabilities. Columns 3 and 4 reveal that using a probit instead of a linear probability model leads to similar conclusions. I conduct several robustness checks in Appendix C.2. There I also show that the same relationship between liquid assets and occupational switching cannot be found for workers leaving manual or abstract occupations.

Current liquid asset holdings and future switching decisions could be jointly determined, and in fact will be so in the quantitative model of the next section. As a result, the estimate of β in equation (1.1) might be biased. I therefore view these estimates only as informative correlations that support the considerations from the small model introduced above. Also, I can use these empirical estimates at a later point to compare them to the results from a synthetic panel, simulated from the quantitative model in Section 5.

[12] Tables 1.5 and 1.6 in Appendix C.2 show results that separate by switches to abstract and manual jobs.
[13] Proxying skill by the log real hourly wage or by raw years of education yields similar results.

Table 1.1: Switching decision and liquid asset holdings.

	(1)	(2)	(3)	(4)
HtM	-3.04***	-2.13*	-0.12***	-0.092**
	(1.17)	(1.16)	(0.045)	(0.045)
Occ. tenure	-1.43***	-1.36***	-0.083***	-0.080***
	(0.098)	(0.094)	(0.0072)	(0.0071)
Skill		8.96***		0.32***
		(1.20)		(0.045)
Controls	Yes	Yes	Yes	Yes
Observations	4904	4904	4904	4904
Model	OLS	OLS	Probit	Probit

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Data are from the PSID, years 1999–2017. Sample selection is as in Cortes (2016), see also Appendix C.2. Dependent variable is whether or not individual leaves routine occ. between t and $t+2$ (coded as 0 or 100). Definition of HtM status is as in Kaplan, Violante, et al. (2014). Skill is a dummy for whether the individual has received more than 12 years of education. Occ. tenure is years of uninterrupted tenure in broad occupation. Additional controls are: region dummies, age, age squared, unionization status, married status, year. Standard errors are clustered at the individual level.

4 Model

Taking into account the empirical findings, this section presents the full model. Time t is continuous and runs forever. The economy consists of a representative firm, heterogeneous households and a government. The representative firm operates under perfect competition and produces using labor from three occupations (manual, routine, abstract) and two types of capital (ICT and non-ICT). Households are subject to uninsurable idiosyncratic labor income risk and die stochastically. The government taxes households and distributes transfers to them. The analysis takes place in general equilibrium, i.e. all factor prices are endogenous. The exogenous driver of technological change is a falling price of ICT capital relative to the consumption good $\frac{1}{q_{ict}}$. I assume that the economy is initially in its steady state in 1980 with a constant q_{ict}. It is then hit by a shock that raises q_{ict} over many years to a new, higher level, until the economy reaches a new steady state. Agents have perfect foresight over the path of q_{ict} once it is revealed in 1980. There is no aggregate risk.

4.1 Representative firm

The final good in the economy Y_t is produced according to a multiply nested constant elasticity of substitution production function (vom Lehn, 2020):

$$Y_t = \underbrace{K_{s,t}^{\alpha}}_{\text{Non-ICT capital}} \left[\mu_m N_{m,t}^{\frac{\gamma_m-1}{\gamma_m}} + (1-\mu_m)\left[\mu_a N_{a,t}^{\frac{\gamma_a-1}{\gamma_a}} + (1-\mu_a) R_t^{\frac{\gamma_a-1}{\gamma_a}}\right]^{\frac{\gamma_a(\gamma_m-1)}{(\gamma_a-1)\gamma_m}}\right]^{\frac{\gamma_m(1-\alpha)}{\gamma_m-1}} \tag{1.2}$$

where

$$R_t = [(1-\mu_r) \underbrace{K_{ict,t}^{\frac{\gamma_r-1}{\gamma_r}}}_{\text{ICT capital}} + \mu_r N_{r,t}^{\frac{\gamma_r-1}{\gamma_r}}]^{\frac{\gamma_r}{(\gamma_r-1)}}$$

Here, $N_{j,t}$ denotes effective labor employed in occupation $j \in \{m, r, a\}$. There are two types of capital, ICT and non-ICT capital (e.g. structures). The laws of motion for the two types of capital are, respectively,

$$\dot{K}_{ict,t} = q_{ict,t} I_{ict,t} - \delta_{ict} K_{ict,t}$$
$$\dot{K}_{s,t} = q_{s,t} I_{s,t} - \delta_s K_{s,t} \ ,$$

where $q_{x,t}$ denotes the amount of capital of type x that can be purchased for one unit of output at time t (Greenwood et al., 1997). The price of the final good is normalized to one. The parameters δ_x capture depreciation of the capital stock. All factor inputs are paid their marginal products and the firm makes zero profits because of the assumption of constant returns to scale. The first order conditions of the firm's profit maximization problem are listed in Appendix D.1.

Crucially, an increase in q_{ict}, i.e. a falling relative price of ICT capital, is the exogenous driving force behind the polarization of labor markets in the model. In line with Eden and Gaggl (2018), I set the relative price of non-ICT capital $\frac{1}{q_{s,t}}$ to one at all times for the remainder of this paper.

4.2 Households

There exists a continuum of mass one of households who value consumption c and leisure $(1-\ell)$. Let $u(c, \ell)$ denote the flow utility function, which is additively separable in consumption and leisure, monotonically increasing in c and monotonically decreasing in ℓ. Households discount the future at rate ρ and die at rate ζ, hence their effective discount rate is $\hat{\rho} \equiv \rho + \zeta$.

In order for the model to generate a share of hand-to-mouth agents that is as high as in the data, I follow Kaplan, Moll, et al. (2018) in modeling two assets that households can save in: a liquid asset m and an illiquid asset $\tilde{k}$. Like them, I assume that households who die are replaced by newborn households holding zero assets.[14] Households can borrow in the

[14]The accidental bequests of dying households are passed on to all living households in proportion to their

liquid asset up to a borrowing constraint $\underline{m}$. The liquid asset pays no interest (representing, for instance, money or low-yielding bonds), but if households borrow they have to pay an intermediation cost $\kappa > 0$.[15] Households can only hold non-negative amounts of the illiquid asset, $\tilde{k} \geq 0$, which pays a return $r_{k,t}$. $\tilde{k}$ is subject to a portfolio adjustment cost $\chi(\tilde{k}, d)$, where positive d denote a deposit and negative d a withdrawal of wealth from the illiquid account.

Labor income Households optimally choose to work in one of the three broad occupations at each instant of time t. Apart from losing occupation-specific human capital, they can costlessly switch between occupations. Denote the occupational choice by $j \in \{m, r, a\}$. Households' (pre-tax) labor income is

$$inc = w_j \cdot \ell \cdot y \ ,$$

where w_j denotes the wage per efficiency unit in occupation j, ℓ labor supply, and y labor productivity.

The log of labor productivity is in turn composed of several terms:

$$\ln(y) = a_j \cdot (\underbrace{s}_{\text{skill}} + \underbrace{\eta}_{\text{shock}}) + \underbrace{h}_{\text{specif. human cap.}} + \underbrace{\epsilon}_{\text{shock}} \tag{1.3}$$

Each household has a skill type s, which is fixed over the lifetime and distributed in the population according to some cumulative distribution function $F(s)$. The skill type is pre-multiplied by an occupation-specific slope parameter a_j, where I assume $0 = a_m < a_r < a_a$. These assumptions are borrowed from Cortes (2016), Jung and Mercenier (2014), and vom Lehn (2020) and imply a comparative advantage of low-(high-) skilled types in the manual (abstract) occupation. This gives rise to an endogenous sorting pattern of skill types into the three occupations.[16]

Figure 1.2, which plots skill on the x-axis and potential earnings in each occupation on the y-axis, visualizes this. For an equilibrium with positive labor supply in each of the occupations to exist, wages w_j must endogenously be ordered as can be seen on the y-axis.[17] The dashed extensions of the solid lines represent hypothetical earnings of households in all three occupations, and the slopes of the lines correspond to the parameters a_j. The cut-offs

current assets.

[15] In Kaplan, Moll, et al. (2018) the liquid asset earns a small interest rate, of 2% annually. I model this asset as non-interest bearing instead as this reduces the number of market clearing factor prices I need to solve for along the transition path from five to four (three wages plus the interest rate on the illiquid asset). To clear the liquid asset market at all times, I assume that the government supply of m is infinitely elastic (see government budget constraint below).

[16] Edin et al. (2021) refer to this structure as a hierarchical Roy model.

[17] For instance, if it were the case that $\ln(w_m) < \ln(w_r)$, no household would choose to work in the manual occupation. To simplify the exposition, the visualized ordering implicitly assumes that $s > 0$, which will not be the case in the calibrated model in which I assume that $s \sim N(0, 1)$. However, in that case, too, a necessary ordering of wages exists for there to be positive labor supply in each occupation.

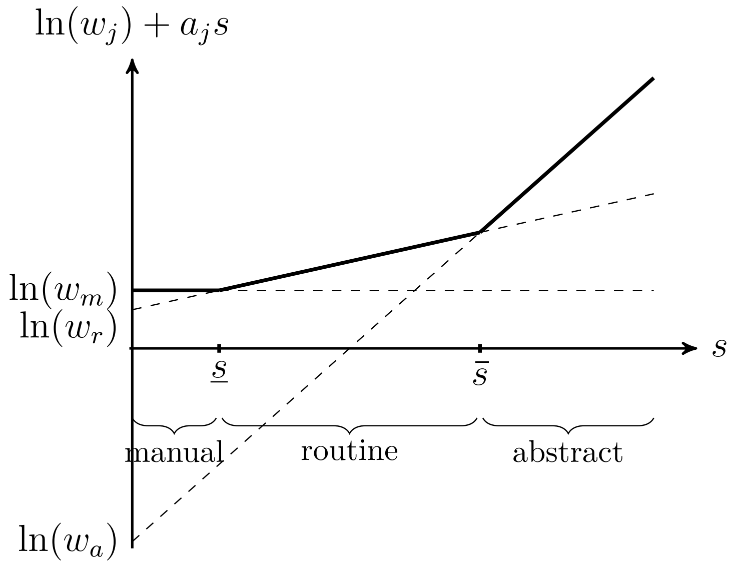


Figure 1.2: Skills and wages before polarization

$\underline{s}$ and $\bar{s}$ separate the skill space into three regions. Absent any other considerations, these cut-offs would sharply divide skill types into working in either of the three occupations in the steady state of the model. As I will discuss further below, I introduce the shock η to capture motives for occupational switches other than technological change.

Figure 1.3 depicts skills and potential earnings after polarization has taken place, i.e. in the new steady state. The exogenous increase in q_{ict} (the fall in the relative price of ICT capital) leads to an endogenous increase in the abstract and the manual wage relative to the routine wage. This shifts the skill cut-offs $\underline{s}$ and $\bar{s}$ inward, leading to a smaller set of skill types who optimally choose the routine occupations in the new steady state.

Hence, there are two margins of adjustment towards the economy with a diminished routine employment share, i.e. when moving from Figure 1.2 to Figure 1.3. First, there is occupational mobility, i.e. net flows of workers switching from the routine to the manual and abstract occupation. Second, some newborn workers choose to work in the manual and abstract instead of the routine occupation, unlike the households they have replaced. Both of these adjustments take place predominantly in the regions $[\underline{s}_{\text{old}}, \underline{s}_{\text{new}}]$ and $[\bar{s}_{\text{new}}, \bar{s}_{\text{old}}]$. I return to an illustrative discussion of the occupational choice over time of workers with skill type $\bar{s}_{\text{old}}$ in Section 5.

Occupation-specific human capital in the current occupation is captured by h in equation (1.3). It can take on two values, capturing whether the worker is inexperienced ($\underline{h}$) or experienced ($\bar{h}^j$), with $\bar{h}^j > \underline{h}\ \forall j$. I allow the size of the spread $\bar{h}^j - \underline{h}$ to differ across occupations in the calibration. For tractability, I assume that households can only be experienced ($\bar{h}$) in their current occupation. Once a household leaves her current occupation and switches to another one, h is set to $\underline{h}$. This implies that only human capital in the current occupation is a state variable, as human capital in the respectively other two occupations is always implicitly $\underline{h}$. It also implies that there is no recall of human capital if a

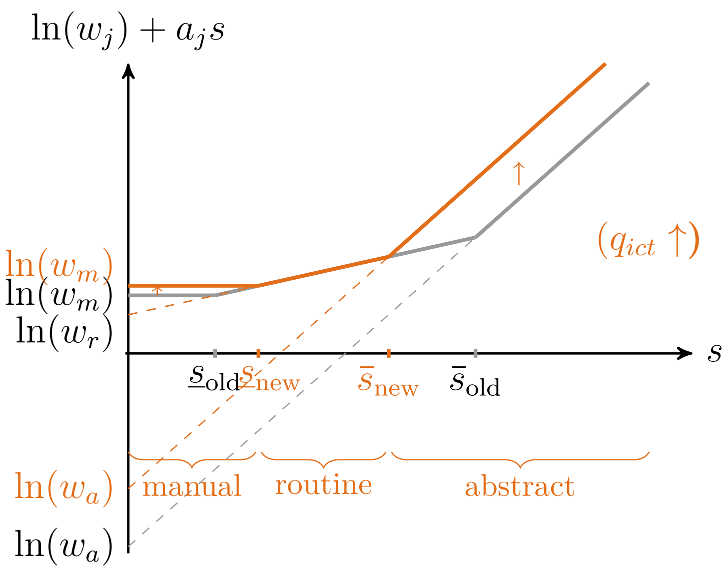


Figure 1.3: Skills and wages after polarization

worker ever returns to one of her previous occupations.[18] Inexperienced households become experienced with occupation-specific Poisson intensity λ_h^j, and experienced households never become inexperienced unless they switch the occupation.

I add two shocks, η and ϵ, to the log of productivity in equation (1.3). η is pre-multiplied by the slope parameters a_j, and hence allows me to introduce occupational mobility in steady state in a parsimonious way. Intuitively, a positive (negative) shock to η shifts households to the right (left) in the schedules of Figures 1.2 and 1.3. A positive shock could be interpreted as a promotion, after which a household who formerly worked in the routine (manual) occupation now works in the abstract (routine) occupation. A negative shock could represent a job loss, after which a household is unable to find another job in the abstract (routine) occupation and therefore takes a routine (manual) job. The advantage of generating steady state occupational mobility in this way is that switches up the occupational ladder (i.e. from manual to routine and from routine to abstract) are contemporaneously correlated with wage gains, and switches down the ladder usually coincide with wage losses. This allows me to target the negative (positive) average wage change of switchers from routine to manual (abstract) documented by Cortes (2016) in the calibration.

The second shock, ϵ, also influences a worker's productivity, but unlike η, not her relative productivity across occupations. I include this shock to generate realistic amounts of earnings risk in the calibration. Each shock evolves according to some stochastic process

$$\dot{\eta}_t = \Phi_\eta(\eta_t),\ \dot{\epsilon}_t = \Phi_\epsilon(\epsilon_t)\ .$$

Newborn households are inexperienced ($\underline{h}$) and start their lives with a draw from the invariant stationary distributions of s, η and ϵ respectively.

[18]This is a common simplification in the literature (Kambourov and Manovskii, 2009a; Kikuchi and Kitao, 2020).

Note that several types of costs resulting from occupational switching are not captured here for the sake of parsimony. In particular, households might not only switch the broad occupation, but at the same time also the industry or the firm, which would entail losses of industry- or firm-specific human capital. There could also be fixed pecuniary costs associated with switching the occupation, such as obtaining a degree or a license, having to move to a new workplace, or buying new work clothes. I view these additional elements as factors that discourage especially those households from switching the occupation that are close to the borrowing constraint. Hence, adding these costs would further increase the relevance of the friction I propose in this paper and potentially make the policies I discuss in Section 6 even more powerful.

Household problem Equation (1.4) shows the household problem.

$$
\begin{aligned}
V_T(i) &= \max_{\{c_t,\ell_t,d_t\},\tau} \mathbb{E}_T \int\limits_{t=T}^{\tau} \mathrm{e}^{-\hat{\rho}(t-T)} u(c_t,\ell_t)\,\mathrm{d}t + \mathrm{e}^{-\hat{\rho}(\tau-T)} \mathbb{E}_T V_\tau^*(i) \qquad (1.4)\\
\text{subj. to:}\quad & \dot{m}_t = (1-\tau_{l,t}) \cdot w_{j,t} \cdot \ell_t \cdot y_t + \mathbf{1}_{\{m_t<0\}} \cdot \kappa \cdot m_t - \chi(d_t,\tilde{k}_t) - d_t + T_t - c_t\\
& \dot{\tilde{k}}_t = r_t \tilde{k}_t + d_t\\
& m_t \geq -\underline{m},\ \tilde{k}_t \geq 0\\
& V_t^*(i) = \max_{\tilde{j}\in\{m,r,a\}} V_t(i_{-\{j,h\}}, \tilde{j}, \underline{h})
\end{aligned}
$$

Here, i collects all idiosyncratic state variables of the households ($\{s,\eta,\epsilon,j,h,m,\tilde{k}\}$), and $\mathbf{1}_{\{.\}}$ denotes the indicator function. T_t denotes a lump-sum transfer from the government, while $\tau_{l,t}$ is a proportional labor income tax. $\tilde{k}_t \equiv \frac{k_{ict,t}}{q_{ict,t}} + k_{s,t}$ denotes capital in units of the final good. The interest rate r_k on $\tilde{k}$, as well as a no-arbitrage condition, which ensures that households are indifferent between holding either type of capital (ICT and non-ICT), are, respectively,

$$
r_t \equiv q_{ict,t} r_{ict,t} - (\delta_{ict} + \dot{q}_{ict,t}/q_{ict,t}) = r_{s,t} - \delta_s\ . \qquad (1.5)
$$

Note also the stopping-time nature of the household problem (1.4): households choose the time $\tau \in [T,\infty)$ at which they switch the occupation. Once they do, the continuation value of the household problem is $V^*(i)$, which is equal to $V(i)$, only that the occupation j is updated and human capital reset to $\underline{h}$, while all remaining state variables ($i_{-\{j,h\}}$) stay unchanged. The household problem gives rise to a Hamilton-Jacobi-Bellman equation, which is shown in Appendix D.2 for the specific processes of the exogenous productivity shocks used in the calibration.

4.3 Government

The government budget constraint holds at each instant of time:

$$G_t + T_t = \tau_{l,t} \int_i inc_i \, \mathrm{d}\Gamma_t(i) + \dot{M}_t^s \; . \tag{1.6}$$

Here, G_t denotes government spending and $\Gamma(i)$ the cumulative distribution function of households over the state space. I assume that fiscal policy always sets M^s such that the liquid asset demand from households is met, i.e. I assume an infinitely elastic supply of M.

4.4 Equilibrium

An equilibrium is defined as paths for household decisions $\{\tilde{k}, m_t, d_t, c_t, \ell_t, j_t\}_{t\geq 0}$, input prices $\{w_{m,t}, w_{r,t}, w_{a,t}, r_{ict,t}, r_{s,t}\}_{t\geq 0}$, government taxes and transfers $\{\tau_{l,t}, T_t\}_{t\geq 0}$, distributions $\{\Gamma_t\}_{t\geq 0}$, and aggregate quantities such that, at every t:

1. Given prices, aggregate quantities, the distribution $\{\Gamma_t\}_{t\geq 0}$, and the stochastic processes for individual states, policy functions c^*, ℓ^*, m^*, $\tilde{k}^*$ and j^* solve the households' problem (1.4).

2. The representative firm optimizes, given input prices. The FOCs (1.10), (1.11), (1.12), (1.13), and (1.14) hold.

3. The government budget constraint (1.6) holds.

4. The labor markets clear, i.e. for $j \in \{m, r, a\}$:

$$N_{j,t} = \int_{i:j^*=j} y_t \ell_t^* \, \mathrm{d}\Gamma_t(i)$$

5. The no-arbitrage condition (1.5) holds and the capital market clears:

$$\frac{K_{ict,t}}{q_{ict,t}} + K_{s,t} = \int_i \tilde{k}_t^* \, \mathrm{d}\Gamma_t(i)$$

6. The liquid asset market clears:

$$M_t^s = \int_i m_t^* \, \mathrm{d}\Gamma_t(i)$$

7. The resource constraint holds:

$$Y_t = C_t + I_{s,t} + I_{ict,t} + G_t + \int_i \chi(\cdot) + \kappa \max\{-m, 0\} \, \mathrm{d}\Gamma_t(i)$$

   where C_t denotes aggregate consumption.

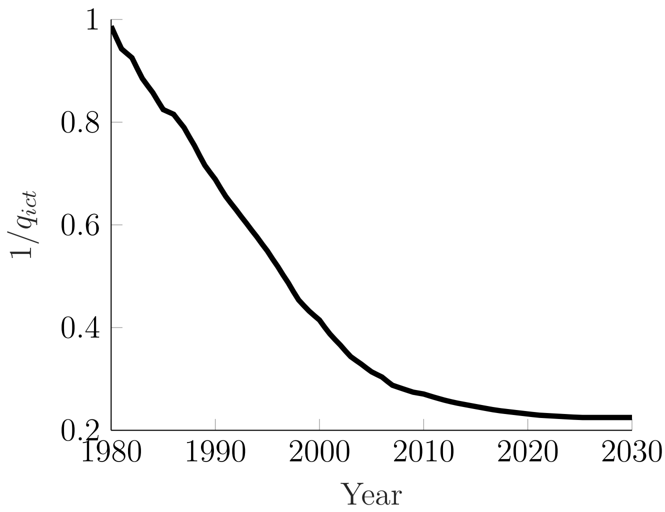


Figure 1.4: Relative price of ICT capital ($\frac{1}{q_{ict}}$).

Notes: Data from 1980 to 2013 are taken from Eden and Gaggl (2018). I assume that $\frac{1}{q_{ict}}$ continues its fall at an average rate of 1% between 2013 and 2025, and stays constant thereafter.

8. The sequence of distributions satisfies aggregate consistency conditions.

4.5 Calibration

Relative price of ICT capital The exogenous driving force of technological change in the model is a fall in the relative price of ICT capital $\frac{1}{q_{ict}}$. To calibrate its evolution over time, I use the estimates reported in Eden and Gaggl (2018), see Figure 1.4. Note that during the last years covered by Eden and Gaggl (2018), i.e. up until 2013, the fall of $\frac{1}{q_{ict}}$ visibly slows down. I assume that the fall continues at the reduced rate of 1% annually between 2014 and 2025 and then stays at its value of 0.225 forever. Agents in the model have perfect foresight over the path of q_{ict}, once it is revealed in 1980.

Externally set parameters I set the first set of parameters as in Kaplan, Moll, et al. (2018). The utility function is

$$u(c, \ell) = \ln(c) - \varphi \frac{\ell^{1+\gamma}}{1+\gamma} ,$$

where γ is set to 1, and φ to 2.2. These choices ensure a Frisch elasticity of labor supply of one and an average labor supply of approximately 0.5.

Households die at rate $\zeta = \frac{1}{180}$, which implies an average life span of 45 years. The unsecured borrowing limit, $\underline{m}$, is set to the average quarterly income. The portfolio adjustment cost function for the illiquid asset $\tilde{k}$ is a convex adjustment cost function

$$\chi(d, \tilde{k}) = \chi_1 \left(\frac{|d|}{\tilde{k}} \right)^{\chi_2} \tilde{k} ,$$

Table 1.2: Externally calibrated values.

Parameter	Value	Description	Source or Target
ζ	$1/(4 \cdot 45)$	death rate	avg. lifetime 45 years
φ	2.2	labor disutility	avg. labor time 8h/day
γ	1	elast. labor supply	Frisch elasticity of 1
$\underline{m}$	avg. qrtly. inc.	borr. constr.	Kaplan, Moll, et al. (2018)
τ_l	0.3	labor tax	Kaplan, Moll, et al. (2018)
T_t	$6\% \cdot Y_t$	lump-sum transfer	Kaplan, Moll, et al. (2018)
α	0.34	Non-ICT cap. share	Eden and Gaggl (2018)
δ_{ict}	0.175/4	deprec. ICT capital	Eden and Gaggl (2018)
δ_s	0.073/4	deprec. struct. capital	Eden and Gaggl (2018)
$[a_m, a_r, a_a]$	$[0, 0.18, 0.77]$	occ. prod. functions	vom Lehn (2020)
$F(s)$	$N(0,1)$	skill distrib.	vom Lehn (2020)

Notes: Rates are expressed as quarterly values. Nominal quantities are expressed in 2017 US dollars.

where χ_1 and χ_2 are parameters.[19] The tax rate on labor income τ_l is set to 30%, and the lump-sum transfer from the government T amounts to 6% of total output Y_t.

The share of non-ICT capital α is set to 0.34 and the depreciation rates of capital to $\delta_{ict} = 0.175$ and $\delta_s = 0.073$ annually, which are average values reported in Eden and Gaggl (2018). For the parameters of the occupational production function I resort to the baseline values used by vom Lehn (2020). Normalizing $a_m = 0$, he estimates $a_r = 0.18$ and $a_a = 0.77$ using data on occupational choices and worker skills, proxied by wages, from the Current Population Survey (CPS). Like him, I assume that skills are standard normally distributed in the population, i.e. $s \sim N(0,1)$. Table 1.2 lists all externally calibrated parameters.

Occupation-specific human capital To calibrate the returns to tenure for each of the three occupations, I follow the methodology outlined in Cortes (2016, Section VI C). In particular, using PSID data from 1981 to 2017, I estimate the following equation:

$$y_{it} = \sum_j D_{ijt} \left(\beta_{j1} Ten_{ijt} + \beta_{j2} Ten_{ijt}^2 + \gamma_{ij} \right) + \delta X_{it} + u_{it}$$

where y_{it} is the log real wage of individual i at time t, D_{ijt} is a dummy indicating whether the individual was working in occupation $j \in \{m, r, a\}$ at time t, Ten_{ijt} is occupational tenure, γ_{ij} is an occupation fixed effect for each individual, and X_{it} are controls (unionization and marital status, region of residence, year dummies and year-occupation dummies). For details regarding sample restrictions and variable definitions, see Appendix C.2.

In line with previous literature (Sullivan, 2010), I find that returns to tenure are higher in the abstract than in the other two occupations (see Figure 1.14 in the appendix). Moreover,

[19] The adjustment cost function in Kaplan, Moll, et al. (2018) contains an additional parameter, χ_0. In a follow-up paper, Alves et al. (2020) use a more parsimonious function, leaving out this additional parameter. I follow this latter approach here.

I do not find statistically significant differences between β_{m1} and β_{r1}, nor between β_{r2} and β_{r2}. I therefore use the estimates of β_{j1} and β_{j2} and approximate the tenure profiles once for the abstract and once jointly for the routine and for the manual occupation. I choose the spreads $\bar{h}^j - \underline{h}$ and the intensities λ_h^j to minimize the mean squared deviations of the return profiles in the model from their empirical analogues over the first twenty years after entering an occupation. Figure 1.14 in the appendix plots these profiles for the calibrated processes, for which I find $\bar{h}^a - \underline{h} = 0.29$ and $\bar{h}^{\{m,r\}} - \underline{h} = 0.15$ as well as $\lambda_h^a = 0.023$ and $\lambda_h^{\{m,r\}} = 0.033$ to provide the best fit. These return profiles are quantitatively in line, as Figure 1.14 also shows, with baseline values from Kambourov and Manovskii (2009b), who estimate returns to occupational tenure unconditionally, i.e. across all occupations.

Productivity shock process Φ_ϵ Following Kaplan, Moll, et al. (2018), I use the productivity shock ϵ to target moments of earnings changes, estimated in Guvenen et al. (2021). I assume that $\Phi_\epsilon(\epsilon_t)$ follows a jump-drift process, with jumps arriving at rate λ_ϵ. At all times, the process drifts toward its mean of zero at rate β_ϵ. Whenever there is a jump, a new log productivity state is drawn from a normal distribution, with $\epsilon_t' \sim N(0, \sigma_\epsilon^2)$. Hence we have

$$\mathrm{d}\epsilon_t = -\beta_\epsilon \epsilon_t + \mathrm{d}J_t \ ,$$

where dJ_t captures the jumps in the process. The targets used for the calibration are shown in Table 1.12 in the appendix. They imply that shocks arrive on average every five years ($\lambda_\epsilon = 0.05$), with a standard deviation σ_ϵ of 1.2 and a half-life of approximately 3 years ($\beta_\epsilon = 0.05$).

Aggregate production function I calibrate the parameters of the aggregate production function (1.2) as is common in the literature (Jaimovich, Saporta-Eksten, et al., 2021; vom Lehn, 2020). In particular, I use the employment shares in the three occupational groups, as well as the share of income accruing to labor, both in 1980 and in 2020, to pin down the six share and elasticity parameters (μ_j, γ_j). For the 1980 employment shares I use the values reported in Autor and Dorn (2013), for the 2020 values I use my own estimates from the 2019 SCF (the most recent employment shares based on CPS data reported in Kikuchi and Kitao (2020) are very similar). For the labor share I take the values reported in Eden and Gaggl (2018) for both 1980 and 2020, assuming that the labor share does not fall further between their latest observation (2013) and 2020. vom Lehn (2020) shows that these six moments together identify the parameters of the production function. As noted above, in accordance with Eden and Gaggl (2018) I assume that the income share accruing to non-ICT capital has been constant over time at $\alpha = 0.34$.

The calibrated values of the production function parameters are as expected (Table 1.3). While routine labor and ICT capital are relatively easy to substitute ($\gamma_r > 1$), abstract labor is relatively complementary to the input provided by both routine labor and ICT

Table 1.3: Internally calibrated values.

Parameter	Value	Description	Target	Data (Model)
Supply side				
μ_m	0.13	PF share man.	1980 Empl. share rout.	58.5% (58.5%)
μ_r	0.94	PF share rout.	1980 Empl. share abstr.	31.6% (31.5%)
μ_a	0.69	PF share abstr.	1980 Labor share	64.0% (64.2%)
γ_m	1.67	PF elast. man.	2020 Empl. share rout.	45.0% (45.4%)
γ_r	2.47	PF elast. rout.	2020 Empl. share abstr.	42.3% (42.5%)
γ_a	0.27	PF elast. abstr.	2020 Labor share	57.0% (56.5%)
Demand side				
$\hat{\rho}$	0.018	discount rate	$\tilde{K}/Y$	2.92 (2.77)
κ	0.034	borr. wedge	M/Y	0.26 (0.30)
χ_1	0.87	portf. adj. cost	share poor HtM	0.10 (0.10)
χ_2	1.35	-	share wealthy HtM	0.20 (0.20)
Wage changes of routine switchers (1980-2020)				
$\eta_h - \eta_l$	0.78	spr. prod. grid	Avg. wage chng. switchers $r \to a$	3.4% (1.5%)
λ_η	0.02	λ prod. shock	Avg. wage chng. switchers $r \to m$	-11.2% (-10.9%)

Notes: Rates are expressed as quarterly values.

capital ($\gamma_a < 1$). The substitution elasticity of manual labor with the nest composed of abstract labor and the routine input is again relatively high ($\gamma_m > 1$).[20] Given the specified production function, the response of wages in each occupation to a rise in q_{ict} in the short term (i.e. when factor inputs are held fixed) can be characterized analytically. Appendix D.3 discusses this and shows that, for the calibrated parameters, manual and abstract wages rise on impact when ICT capital becomes cheaper and routine wages fall.

Demand side Concerning the demand side of the economy, I need to calibrate four parameters that are also present in Kaplan, Moll, et al. (2018). I use the same targets as they do to pin them down. In particular, I calibrate the effective discount rate $\hat{\rho}$, the borrowing wedge κ and the parameters of the portfolio adjustment cost function $\chi(\cdot)$ using both the ratio of liquid and illiquid assets to output in the economy, and the shares of poor and wealthy hand-to-mouth households. I target these four statistics in the initial steady state of the model. I find very similar values for these five parameters as Kaplan, Moll, et al. (2018), and hit the targets relatively well. Only the calibrated value of κ, which implies an annual borrowing wedge of 13.7%, is notably higher than in Kaplan, Moll, et al. (2018) (6.0%), but still well in line with empirical estimates of the price for unsecured borrowing (Dempsey and Ionescu, 2021).

[20] The calibrated values of these parameters are similar to those used in the previous literature. For instance, Jaimovich, Saporta-Eksten, et al. (2021) calibrate $\gamma_m = 1$, $\gamma_r = 1.89$ and $\gamma_a = 0.32$. vom Lehn (2020) uses equipment instead of ICT capital in the production nest with routine labor, rendering his production function parameters less comparable to the ones I use here.

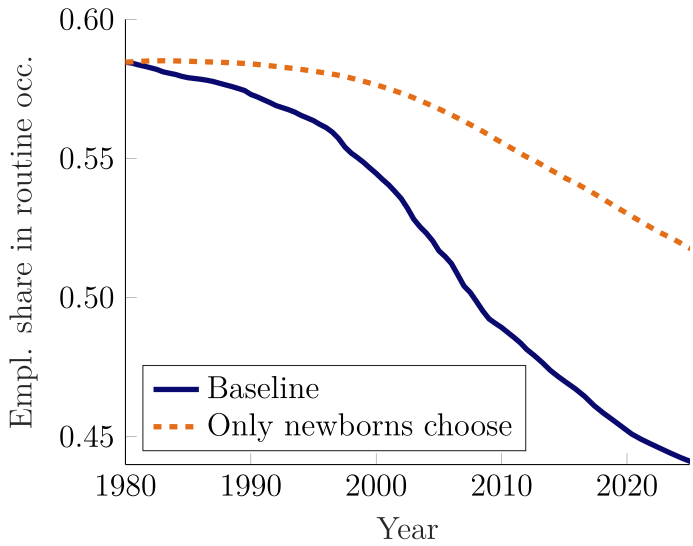


Figure 1.5: Employment share in the routine occupation.

Notes: Blue solid: baseline transition. Orange dashed: counterfactual with initial occupational choices made permanent for rest of life.

Productivity shock process Φ_η Lastly, I use the productivity process Φ_η to target the initial wage change of workers who leave the routine occupations relative to those who stay, estimated by Cortes (2016). To be as parsimonious as possible, I assume that η follows a two-state Poisson process, with symmetric transition rate λ_η between the states. This yields two parameters to target the two statistics: the spread between the states, $\eta_h - \eta_l$, and the transition rate λ_η. To obtain statistics that are comparable to Cortes (2016), I simulate a panel of 10,000 households along the transition path between 1980 and 2020 and perform the same regressions that he uses to produce his estimates. The calibrated shock hits the workers on average every twelve years and the impact of a standard deviation shock on log productivity is significantly smaller than that of a typical shock to the other productivity variable, ϵ.

4.6 Untargeted moments

The role of newborns in the decline of the routine employment share Cortes, Jaimovich, Nekarda, et al. (2020) provide empirical evidence highlighting that 34%–43% of the fall in routine employment has been due to a reduced propensity to enter these occupations from non-employment and unemployment. They further show that the declining propensity of young workers (16–34 years) has been especially important in accounting for the drop in aggregate routine labor. While I do not model unemployment, one can interpret the entry of newborn households as entry from non-employment. I therefore conduct a counterfactual exercise to assess whether my model performs realistically in this regard.

Starting in the initial steady state, I use the optimal occupational policy functions j_t^* to decide which occupation new labor market entrants choose at the beginning of their life. Afterwards, I counterfactually assume that they never switch the occupation, i.e. I iterate forward the distribution of households over the state space assuming that households only make an occupational choice at the beginning of their lives. Hence, this is an out-of-equilibrium exercise that abstracts from general equilibrium effects on initial occupational choices and on wages, just as in Cortes, Jaimovich, Nekarda, et al. (2020). Details on how I construct this and all following counterfactuals are relegated to Appendix D.4.

The orange dashed line in Figure 1.5 depicts the resulting mass of routine workers in the economy. It falls much more slowly compared to the baseline transition (blue solid), which implies that some of the fall in routine labor in the model is due to occupational switching. Consistent with Cortes, Jaimovich, Nekarda, et al. (2020), in 2020 41% of the fall in routine employment since 1980 is accounted for in the model by the reduced propensities of new labor market participants to enter the routine occupations.

Relative wages and interest rate While I targeted the changes in employment shares in the calibration, the evolution of relative wages per efficiency unit w_a/w_r and w_m/w_r over time were left untargeted. Cortes (2016) estimates these and finds that the log of w_a/w_r rose by about 25% between 1980 and 2007, and the log of w_m/w_r by 10 to 15%.[21] The model produces a nearly identical rise in w_a/w_r, though only a somewhat smaller increase of up to 4% in w_m/w_r (see Figure 1.23 in the appendix). Also, while in the data relative wages in the abstract and manual occupations grow quite fast in the 1990s, in the model the growth takes place more slowly, and hence the evolution of w_a/w_r lags behind that in the data by about 10 to 15 years.[22]

Figure 1.24 in the appendix further plots the evolution of the interest rate. Since firms demand more capital when its relative price falls, the interest rate rises over time. This is in line with empirical evidence discussed in Moll et al. (2021).

Wage changes upon occupational switch In the calibration I targeted the average year-on-year wage change upon leaving the routine for either the manual or the abstract occupation, compared to those who stayed in the routine occupation (-11.2% for exits to manual, 3.4% for exits to abstract). As discussed, Cortes (2016, Table 3) also compares wage changes over longer horizons, documenting faster wage growth for switchers than for stayers. I visualize his estimates and the corresponding statistics from the simulated panel of my model in Figure 1.13 in the appendix. Overall, I am able to capture the differential wage paths of switchers compared to stayers quite well, though for both occupations long-run wage growth is a bit less pronounced in the model than in the data.

[21] These numbers correspond to the lower left panel of Cortes (2016, Figure 6), as in that specification he allows for occupation-specific tenure profiles, as I do in my model.

[22] All these statistics refer to wages per efficiency unit across occupations (w_j). Concerning average wages ($w_j \cdot y$), once aggregated to the three broad groups used in this paper, Autor and Dorn (2013) report a difference between mean log hourly wages in the abstract and in the routine occupations of 0.30, and of 0.43 between the routine and the manual occupations in 1980. The corresponding statistics in the model are 0.34 and 0.17 respectively.

Hand-to-mouth shares by occupation While the calibration targeted unconditional shares of hand-to-mouth households in the economy, it left the shares conditional on occupation untargeted. These moments are reported in Table 1.13 in the appendix for the initial steady state of the model as well as for the earliest available wave of the SCF (1989). Overall, the model provides a very good fit to the data in this regard.

5 Transition Path

This section studies the transition between the two steady states. The focus lies on demonstrating that the existence of a large share of households who hold few liquid assets, i.e. who are close to the borrowing constraint, causes a lag in the reallocation of labor into the abstract and manual occupation.

5.1 Individual decision rules

To exemplify this point, I first analyze the switching decisions of a particular set of formerly routine workers whose optimal occupational choice is clearly affected by technological change. Specifically, I consider the occupational policy functions j_t^* of experienced ($\bar{h}$) routine workers of skill type $s = \bar{s}_{\text{old}}$, and ask at what point in time they decide to exit the routine for the abstract occupations. Workers of this skill type predominantly work in the routine occupations before and in the abstract occupations after polarization has shifted relative wages in favor of performing abstract work (cf. Figure 1.3).

Figure 1.6 visualizes this choice over time, once for the households that are hand-to-mouth, and once for those that are not. It plots the binary choice x_t of entering the abstract (and hence leaving the routine) occupation, averaged across wealth (m and $\tilde{k}$), conditioning on exogenous productivity states $\eta = \eta_h$ and $\epsilon = \mathbb{E}[\epsilon] = 0$, evaluated using the stationary distribution of households in the initial steady state $\Gamma(i)$, i.e.[23]

$$x_t = \frac{\int_{i:\mathcal{I}} \mathbf{1}_{\{j_t^*=a\}} \mathrm{d}\Gamma(i)}{\int_{i:\mathcal{I}} \mathrm{d}\Gamma(i)} \ , \tag{1.7}$$

where $i : \mathcal{I}$ implies that individual i belongs to the set $\mathcal{I} = \{s = \bar{s}_{\text{old}}, j_{1980}^* = r, h = \bar{h}, \eta = \eta_h, \epsilon = 0\}$. For the orange dashed line in Figure 1.6, corresponding to choices of hand-to-mouth households I replace $\Gamma(i)$ in (1.7) with

$$\Gamma(i) \cdot \mathbf{1}_{i:\{m=0 \vee m=-\underline{m}\}} \ ,$$

[23]Using the current distribution of households over the state space at each point in time $\Gamma_t(i)$ in the computation of x_t leads to very similar results. However, as experienced routine workers of this skill type keep leaving the occupation (or die), the mass of this type of workers approaches zero over time. Computing x_t over an increasingly small set of workers leads to large jumps in x_t during the later years of the transition. Therefore, I display the exit probabilities that use the (constant) steady state distribution at all times here.

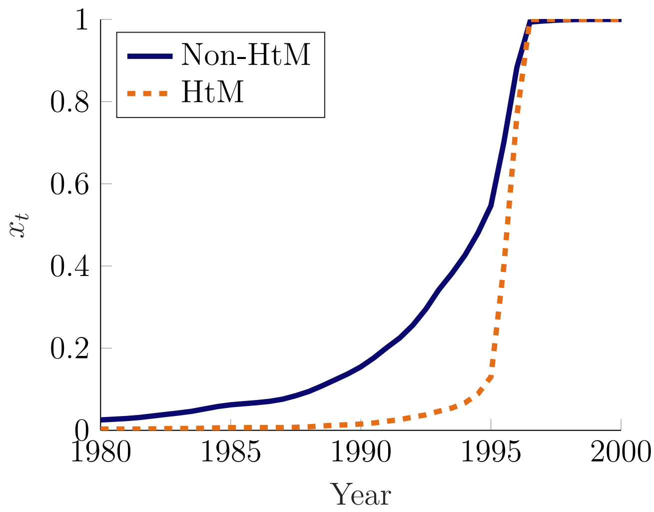


Figure 1.6: Mass of experienced rout. workers with $s = \bar{s}_{\text{old}}$ who switch to the abstr. occ.

Notes: Binary choice x_t of entering the abstract occupation, averaged across wealth (m and $\tilde{k}$), conditioning on exogenous productivity states $\eta = \eta_h$ and $\epsilon = \mathbb{E}[\epsilon] = 0$, and evaluated using the stationary distribution of households in the initial steady state $\Gamma(i)$. The dashed red line depicts the case when averaging over individuals with $m = 0$ or $m = -\underline{m}$, the solid blue line when averaging over all remaining individuals.

and I replace it with

$$\Gamma(i) \cdot \mathbf{1}_{i:\{m \neq 0 \wedge m \neq -\underline{m}\}}$$

for the blue solid line that corresponds to occupational choices of non-hand-to-mouth households. A value of $x_t = 0$ indicates that the mass of workers from the set $\mathcal{I}$ who would leave the routine for the abstract occupations at time t equals zero, while $x_t = 1$ implies that all of these workers would make the switch if they found themselves working in the routine occupation at time t.

Figure 1.6 can be interpreted as follows. As the relative wage in the abstract relative to the routine occupation rises over time, an increasing share of experienced routine workers with skill type $s = \bar{s}_{\text{old}}$ finds it optimal to switch the occupation. However, while the share of non-hand-to-mouth workers wanting to switch rises quickly, almost none of the hand-to-mouth households prefer a switch up until the mid-1990s. This shows how being close to the borrowing constraint delays investment into future earnings growth, just as it did in the simplified model of Section 2.

5.2 Aggregate implications

Figure 1.6 has shown the differential switching behavior of workers who are at a particular, exemplary, point of the state space. To assess whether the distortions at the individual level have had an impact on aggregate employment shares in the abstract and manual occupations, I simulate a counterfactual transition in which all households choose their occupation like those that are rich in liquid assets and hence far away from the borrowing constraint.

Specifically, I start with the actual employment share in occupation j in the initial

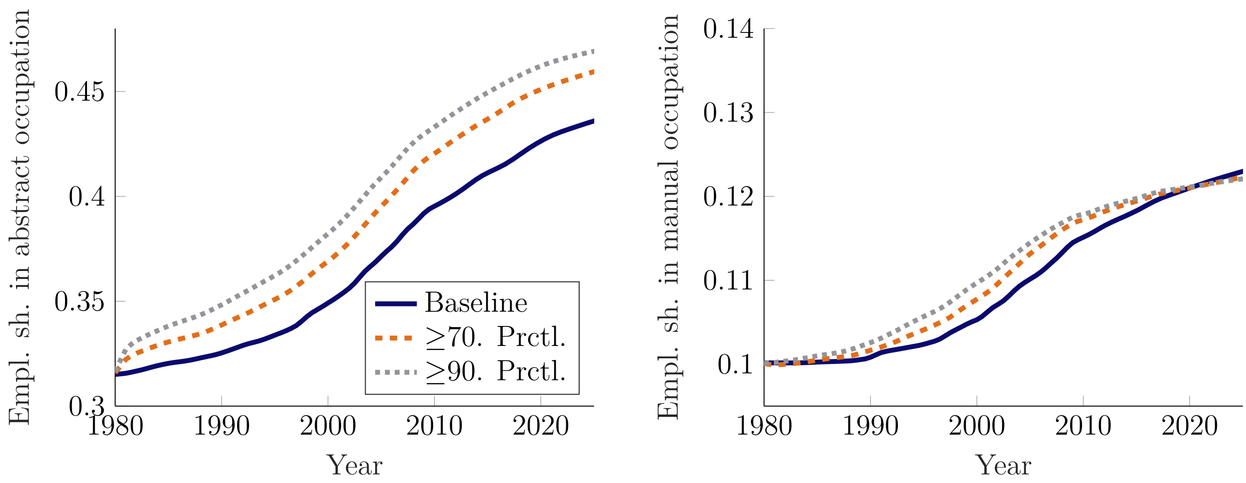


Figure 1.7: Counterfactual empl. share in abstract (left) and manual (right) occ.
Notes: "≥70. Prctl." ("≥90. Prctl.") corresponds to counterfactual simulation in which all households choose occupations like households with m_t greater or equal to the 70th (90th) percentile of the liquid wealth distribution.

steady state. I then iterate this share forward in time using the following procedure. For a given point in time t, I consider all households with liquid assets m_t exceeding the 70th (90th) percentile of the liquid wealth distribution. I then ask what share of these households optimally chooses to work in occupation j, using the policy functions from the baseline transition. I use this probability to iterate forward the employment share and then redo the procedure for the next point in time.[24] At each step, I correct for the fact that liquidity-rich households systematically differ in terms of exogenous productivity (η and ϵ) from the unconditional distribution by reweighting accordingly. Appendix D.4 explains the details. Note that this is not an equilibrium exercise, as I neither specify where the additional liquid assets come from, nor do I let aggregate quantities respond to the changed behavior of households.

Figure 1.7 shows the results. For both the abstract as well as the manual occupation it becomes apparent that had all households acted as if they were far away from the borrowing constraint, the employment shares during the transition would have been significantly higher. For instance, between 1980 and 2020 the employment share in the abstract occupation would have been on average two (three) percentage points higher than in the baseline transition. The effect is also present for the manual employment share, though overall movements are much smaller.[25]

Another way to address whether individual behavior, as depicted in Figure 1.6, becomes

[24]When solving for the transition path I discretize time into periods (Kaplan, Moll, et al., 2018).

[25]The counterfactual abstract employment shares stay elevated for several decades and then slowly approach the baseline. They stay elevated by about 0.5 percentage points even far into the future, however. This is because of the occupation-specific tenure profiles, which are relatively steep in the abstract occupation. This leads to some households inefficiently working in the routine occupation if they do not have enough liquid assets to make the investment of starting a career in the abstract occupation which would promise higher wages only in the future. See the discussion below on introducing the policies in the steady state of the model.

Table 1.4: Switching decision, linear probability and probit models.

	(1)	(2)	(3)	(4)
HtM	-0.32***	-0.44***	-0.062***	-0.12***
	(0.063)	(0.062)	(0.012)	(0.013)
Occ. tenure	-4.43***	-4.02***	-0.60***	-0.59***
	(0.12)	(0.12)	(0.012)	(0.012)
Skill		4.98***		1.04***
		(0.093)		(0.021)
Time trend	Yes	Yes	Yes	Yes
Observations	219132	219132	219132	219132
Model	OLS	OLS	Probit	Probit

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Data are from a synthetic panel of 10,000 households, simulated from the model at annual frequency between 1980 and 2020. Dependent variable is whether or not individual leaves routine occ. between t and $t + 2$ (coded as 0 or 100). Occ. tenure is a dummy for whether or not worker is experienced.

visible at the aggregate level, I rerun the regressions conducted at the end of Section 3 (equation (1.1)). This time, instead of using data from the PSID, I use the synthetic panel of 10,000 households described in the calibration. Mimicking the empirical implementation, I regress a dummy for whether a household switches between year t and $t + 2$ on a dummy for whether a household is hand-to-mouth, controlling for tenure in the broad occupation (proxied by a dummy indicating whether an individual is experienced or not), a linear time trend and in some specifications also for skill s. The results are shown in Table 1.4. Reassuringly, the same negative association between current liquid asset holdings and future switching decision that I found in the data also prevails in the model. In terms of magnitudes, the point estimates are smaller (in absolute terms) than their empirical counterparts in Table 1.1, but all statistically significant.

6 Policy Analysis

The fundamental cause of the protracted labor market transition highlighted so far is the borrowing constraint. It causes households with few liquid assets to underinvest in future earnings growth and to stay in the declining routine occupation for a relatively long time. In this section, I ask what the policymaker can do in order to alleviate this friction. In particular, I study the consequences for social welfare and output of two policies, which I introduce into the model one at a time. The first is a government loan program, the second a transfer—or wage replacement—program. Both policies aim to alleviate the borrowing constraint for workers leaving the routine occupation.

6.1 Design of the policies

The policies are supposed to target those routine workers for whom switching to the manual or abstract occupation is costly in the short run, but profitable in the long run. Therefore, ideally, I would like to condition on households experiencing a wage loss upon switching. This would, however, require making past wages an additional state variable in the model, which would greatly increase the computational burden of solving it. Instead, I proxy for this condition in two ways. First, I only entitle experienced ($\bar{h}$) routine workers who leave the occupation to the programs, as they are the ones who are most likely to suffer from temporary earnings losses due to the depreciation of occupation-specific human capital. Second, to avoid offering the program to households for whom a switch to the abstract occupation reflects career progression, I limit eligibility to those skill types (s) among which the majority of households is currently employed in the routine occupation.[26]

All households begin their lives as non-eligible, and only experienced routine workers who switch to the abstract or manual occupation become eligible, provided that their skill type is qualified to receive the treatment. Households are never restricted in their occupational choice after they have selected into treatment, i.e. they can move back to the routine occupation at a later point if they find this optimal. Note, however, that because previous human capital cannot be recalled, they are inexperienced ($\underline{h}$) routine workers if they do so.

For the loan program, I assume that eligible workers receive a loan of size L in liquid assets from the government upon leaving the routine occupation. After having received the loan, they pay it back at a constant rate $\zeta \cdot L$ until they die. This ensures that in expectation the government breaks even on each individual loan, even though on aggregate it might face temporary deficits (surpluses) from the program if at a given time many workers are taking out (paying back) a loan.

The wage replacement program aims to replace part of the lost earnings that arise due to the occupational switch, and I model it as a transfer payment to the eligible households. This policy is motivated by the Reemployment Trade Adjustment Assistance (RTAA) program currently in place in the U.S., under which wage replacements are paid to workers who lost their job because of foreign trade, and found reemployment at a lower wage than before. In the model, I assume that once a worker becomes eligible (i.e. once an experienced routine worker leaves the occupation), she collects a type-independent monthly benefit of R. Workers lose eligibility, and hence leave the program, at a quarterly rate of λ_I. I choose λ_I to coincide with the average quarterly rate at which inexperienced workers become experienced in the economy of 0.03, which leads to an average duration in the program of about eight years.

[26]This condition rules out relatively high-skilled households from becoming eligible. These are households who only work in the routine occupation (if at all) in case they receive a low productivity draw η_l. Once they receive a high draw (e.g. a promotion), they switch to the abstract occupation, usually incurring a wage gain. While this pattern in principle holds for every skill type, the further to the right of the skill distribution, the higher the comparative advantage from working in the abstract occupation and hence the higher the average wage gain from switching there (Figure 1.3). I do not include such a condition on skill type for switchers to the manual occupation, as here wage gains occur much less often.

There are two key differences in the design of the wage replacement and the loan program. First, workers do not pay back the wage replacement payments they have received. Second, in the spirit of the RTAA program, payments are conditioned on working in the new (abstract or manual) occupation. If a worker decides to switch back to the routine occupation at some point, she stays entitled but does not receive the benefit R unless she returns to the manual or abstract occupation.

To balance its budget, the government adjusts the labor income tax $\tau_{l,t}$, while government expenditures G_t and lump-sum transfers to the households T_t stay unchanged compared to the baseline transition. I assume that the policies start in 1980 and that workers can enter the programs until 2025. All payments are stopped in 2070, an average lifetime after the last workers were allowed into the program.[27] Finally, I use the following welfare criterion to evaluate the policies:

$$\Delta W = \int\limits_{t=1980}^{\infty} \int\limits_{i} \left[u(c_t^{pol}, \ell_t^{pol})\, \mathrm{d}\Gamma_t^{pol}(i) - u(c_t^{basel}, \ell_t^{basel})\, \mathrm{d}\Gamma_t^{basel}(i) \right] \mathrm{d}t$$

This compares (unweighted) utilities of all individuals alive at time t under the policy and in the baseline (inner integral), and then sums this difference over time (outer integral). This ensures that the utility of later cohorts (e.g. of those joining the labor market in 2000 or 2020) is not discounted relative to those alive in earlier years.

6.2 Results

I now study the results of introducing the two policies in turn. I separately consider making the policy only available to switchers to the abstract and only to switchers to the manual occupation, as I find that the implications of the policies differ by the direction of switch. Since the increase in employment in the abstract occupation has been quantitatively much more important than that in the manual occupation since the 1980s, I focus here on implementing the program solely for workers who switch to the abstract occupation. Targeting the policies at exits to the manual occupation is relegated to Appendix E.

Welfare comparison

The blue solid lines in Figure 1.8 show the absolute difference between the welfare criterion under the policy and under the baseline for different values of L (left panel) and R (right panel). In both cases, welfare differences are inversely u-shaped in the size of the policy, with the optimal loan being 10,000$ and the optimal monthly wage replacement 420$.[28]

[27] I phase out the programs so that the economy transitions back to the same steady state under the policy as in the baseline (which features no policy).

[28] Conditional on observing a year-on-year earnings loss when switching from the routine to the abstract occupation, the median monthly earnings loss in the synthetic panel simulated from the model is 680$. Hence, the optimal wage replacement program replaces about 60% of the median earnings loss.

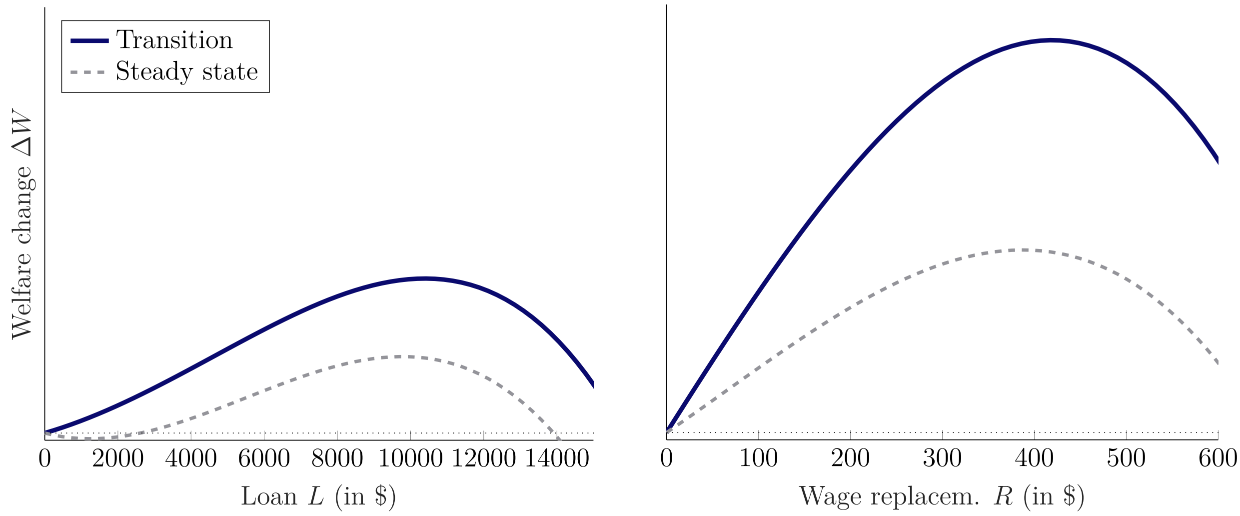


Figure 1.8: Absolute changes in welfare under the policies compared to the baseline.
Notes: “Transition”: baseline transition path including technological change ($q_{ict,t}$ rises). “Steady state”: transition during which no technological change occurs ($q_{ict,t}$ stays constant at its 1980 level). Nominal quantities are expressed in 2017 US dollars.

Furthermore, the optimal wage replacement leads to a larger increase in welfare than the optimal loan.

There are several reasons why the policies can increase welfare vis-à-vis the baseline. First, they loosen the borrowing constraint for eligible households and therefore improve their ability to smooth consumption. In doing so, they enable these households to invest in future earnings growth, given that they decide to stay in the abstract occupation in the long run. Second, at least in the case of the wage replacements, income is redistributed from tax payers to the recipients of the transfers. Third, there are general equilibrium effects. Output potentially grows due to an improved allocation of labor across occupations. Also, the fact that more routine workers find it optimal to switch to the abstract occupation when the policies are in place puts upward (downward) pressure on the routine (abstract) wage. This is illustrated in Figure 1.9, which plots wages in the three broad occupations both along the baseline transition (solid lines) and along the transition that implements the optimal wage replacement program (dashed lines). These wage changes cause incomes to rise for medium-skilled routine workers and vice versa for abstract workers, which leads to aggregate welfare gains due to the strictly concave utility function. Figure 1.9 also reveals that, as some low-skilled workers switch from the manual to the routine occupation because of the now higher wages there, manual wages rise as well due to lower supply of manual labor. This further increases welfare as it benefits those low-skilled workers who remain in the manual occupation.

There are, however, also negative effects of the policies, which outweigh once the loan (wage replacement) becomes very large. First, both policies have to be financed by distortionary labor taxes. Even in the case of the loan program, where the government breaks even in expectation on each individual loan, it has to raise taxes whenever the paybacks

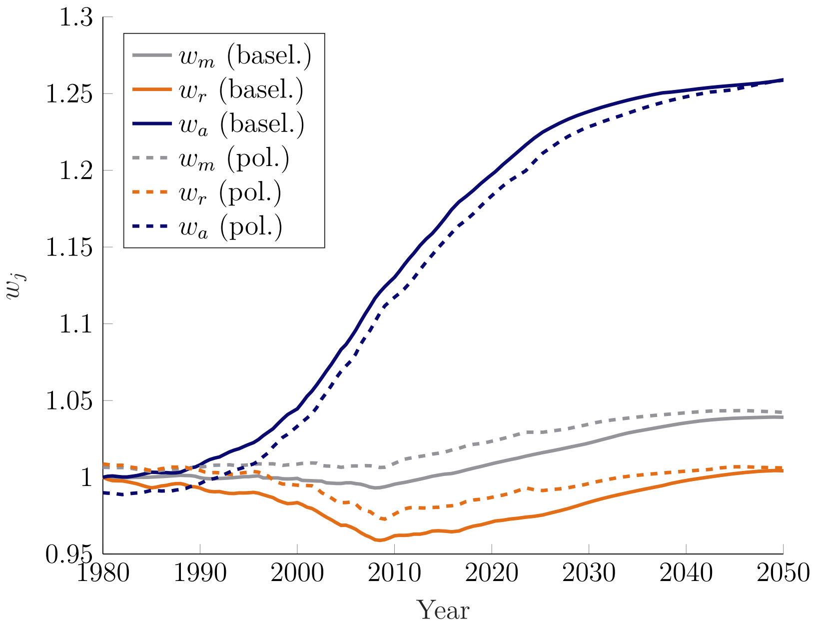


Figure 1.9: Wages per efficiency unit of labor.

Notes: Solid lines: Baseline transition. Dashed lines: Transition with optimal wage replacement program.

from earlier loan recipients are not high enough to cover the initiation of new loans.[29] Second, it is important to understand that every time an experienced routine worker switches to the abstract occupation, occupation-specific human capital gets destroyed. Hence, every additional switch induced by the policies entails a resource cost. For some workers it might be efficient to incur this short-term cost, as switching to the abstract occupation can lead to higher income in the future (Section 2). However, some households might opt into the program by switching into the abstract occupation only in order to pick up the loan (wage replacement) and be able to smooth current consumption, without any intent of staying in the abstract occupation in the medium to long run.

This also helps explain why, by design, the loan program obtains smaller welfare gains than the wage replacement program. Under the loan program there is no disadvantage for the individual worker to switching back to the routine occupation after having obtained the loan. Hence there exist many experienced routine workers who, when hit by an adverse income shock, opt into the program to collect the loan and then switch back to the routine occupation relatively quickly. While this can raise (short-term) utility of the individual worker, once the loaned money has been spent the worker is left with lower human capital and hence lower medium- to long-run earnings than if the policy had not been offered.[30] In contrast to this, under the wage replacement program the payments are conditioned on working in the abstract occupation (in the spirit of the RTAA program). Hence, the workers

[29]The labor taxes needed to finance the optimal programs are shown in Figure 1.26 in the Appendix.

[30]The fact that some experienced workers in the model leave their occupation only for an infinitesimal short period of time and still return as an inexperienced worker is clearly an abstraction that owes to the assumption of no recall of previous human capital. This assumption is necessary to keep the model computationally tractable. More realistically, labor market frictions might prevent households from switching back and forth between occupations at great speed, causing a gradual depreciation of occupation-specific human capital.

who opt into this program are predominantly those who have an interest in becoming abstract workers and leaving the routine occupation for good.

In short, the wage replacement program is more successful in incentivizing the right workers to leave the routine occupation than the loan program. To demonstrate this effect, I compute the share of program participants who have returned to working in the routine occupation among all currently entitled households at time t. Averaged across the years 1980 to 2025 this share is 33% under the optimal loan program and only 11% under the wage replacement (see Figure 1.25 in the appendix). This indicates a significant resource cost of the loan program, caused by workers who switch to the abstract occupation despite having no intent of working there in the long run.

Welfare change in steady state

The grey dashed lines in Figure 1.8 show the welfare implications of introducing the policies in the initial steady state of the model.[31] About half of the welfare gains obtained along the actual transition path are also realized in the steady state. This has two main reasons. First, the same qualitative effect of the policies on wages depicted in Figure 1.9 are also present in the steady state, leading to a shift of income from rich abstract workers to less rich routine and manual workers. Second, even though wages are constant in the steady state, returns to tenure ($\bar{h}^j - \underline{h}$) are higher and hence earnings paths steeper in the abstract than in the routine occupation. Therefore, even in the steady state the borrowing constraint prevents some workers from optimally borrowing against high future income and choosing the abstract instead of the routine occupation. Figure 1.8 demonstrates that these two reasons are sufficient to cause a rise in welfare when the policies are introduced in the steady state, despite the destruction of occupation-specific human capital the additional switching entails. Nonetheless, the welfare gains are only half of those obtained when implementing the policies along the transition path, and, as I will return to below, the policies' effect on output is negative in steady state, while it is positive on average along the transition path with technological change.

Welfare over time

How are the welfare gains from implementing the optimal loan (wage replacement) program distributed over time? To answer this question, Figure 1.10 plots the change in welfare under the respective policies compared to the baseline in terms of consumption equivalent units. More specifically, it shows ϕ_t which satisfies:

$$\int_i u(\phi_t \cdot c_t^{basel}, \ell_t^{basel}) \, \mathrm{d}\Gamma_t^{basel}(i) = \int_i u(c_t^{pol}, \ell_t^{pol}) \, \mathrm{d}\Gamma_t^{pol}(i) \; .$$

[31] Technically, I still simulate a transition to compute the welfare criterion, but one in which q_{ict} stays constant at its 1980 level. I then phase out the policies between 2025 and 2070, which makes the results comparable to the policy analysis along the actual transition path with a falling relative price of ICT capital.

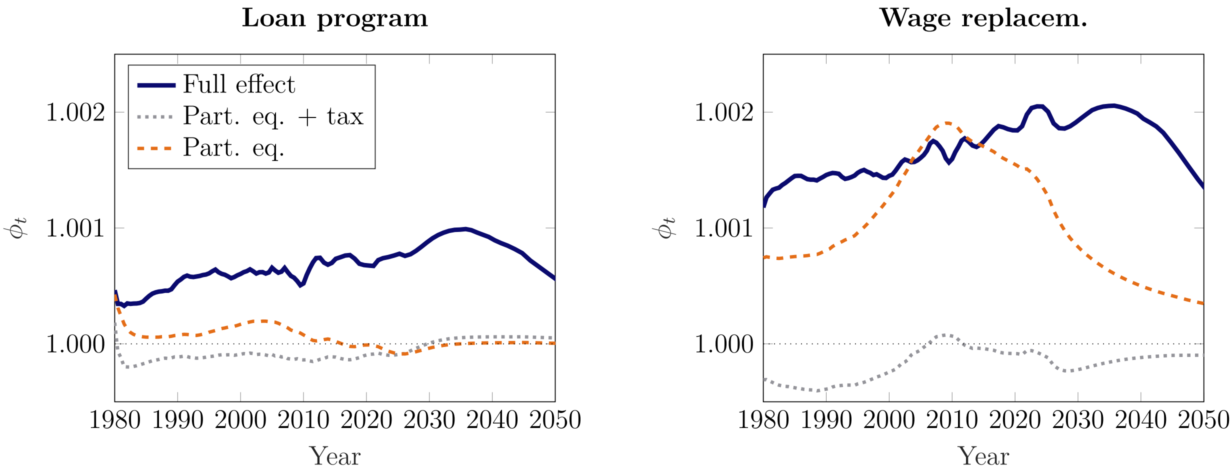


Figure 1.10: Changes in welfare under optimal policies compared to the baseline over time. *Notes:* Left: optimal loan program. Right: optimal wage replacement. ϕ_t is s.t. $\int_i u(\phi_t \cdot c_t^{basel}, \ell_t^{basel})\, \mathrm{d}\Gamma_t^{basel}(i) = \int_i u(c_t^{pol}, \ell_t^{pol})\, \mathrm{d}\Gamma_t^{pol}(i)$ holds. "Part. eq." holds fixed taxes $\tau_{l,t}$ and factor prices at their values from the baseline transition without policy. "Part. eq. + tax" holds fixed factor prices only. "Full effect" is the full general equilibrium effect.

The blue solid line in Figure 1.10 shows that both programs lead to welfare improvements along the entire transition path. In terms of magnitudes, every household in the economy could increase consumption by up to 0.1% (0.2%) under the optimal loan (wage replacement) program during the later stages of the transition.

To better understand what is driving these welfare improvements, I further dissect the total general equilibrium effect of the policy. The orange dashed line in Figure 1.10 plots ϕ_t when keeping the labor tax as well as all factor prices, i.e. wages and the interest rate, constant at their values from the baseline transition without policy. The only change compared to the baseline is that the policies are implemented (without being funded). As can be seen, this partial equilibrium effect raises welfare at most (all) horizons under the loan (wage replacement) program. Households benefit from the loans (wage replacement payments) and increase consumption and leisure. Only at later stages of the loan program, when no new households become eligible and earlier recipients have to pay back the loan they received, the welfare effect turns slightly negative. On top of this partial equilibrium effect, the grey dotted line further lets the labor income tax $\tau_{l,t}$ adjust to the level needed to finance the payments. This takes away from the welfare gains obtained initially, and the overall effect becomes slightly negative at most horizons. Finally, the blue solid line plots the full general equilibrium effect, i.e. after wages and the interest rate have adjusted to clear the factor markets. This effect is positive and quantitatively the most important one at almost all horizons of both programs.

The total effect of the policies on welfare, however, masks substantial heterogeneity both across birth cohorts and especially across skill types. To demonstrate this, I consider welfare changes in terms of expected lifetime consumption of newborn workers next. This is depicted in Figure 1.11 for the optimal loan program (left panel) and the optimal wage replacement (right panel). The graphs plot differences in expected lifetime consumption of newborns

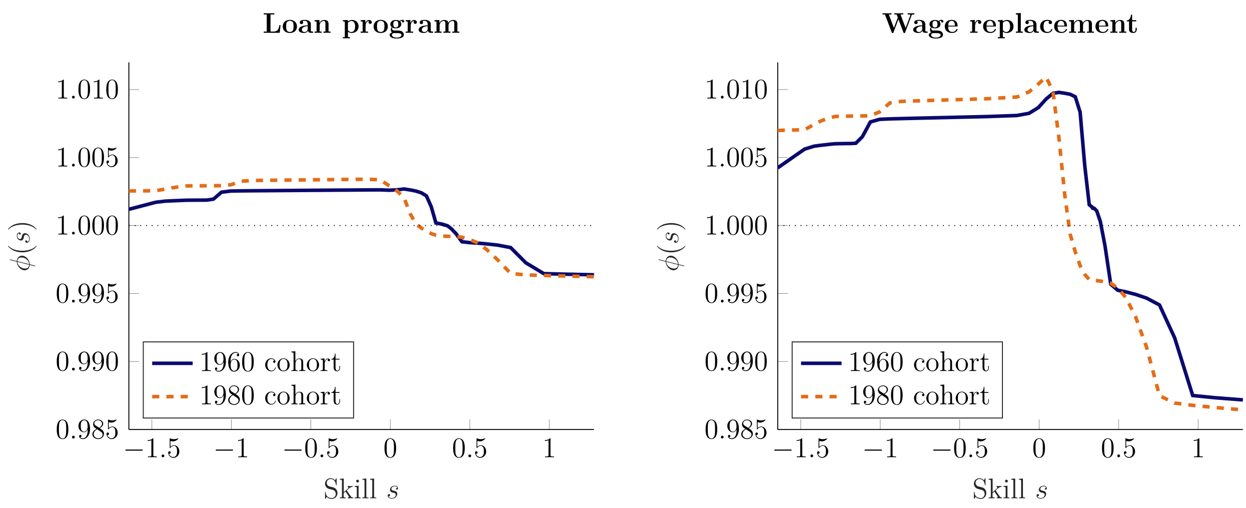


Figure 1.11: Changes in expected lifetime consumption of newborn households under optimal policies compared to the baseline.

Notes: Left: optimal loan program. Right: optimal wage replacement. $\phi(s)$ is such that $\mathbb{E}_{\tilde{T}}\left[\int_{t=\tilde{T}}^{\infty} \mathrm{e}^{-\hat{\rho}(t-\tilde{T})}u(\phi(s)\cdot c_t^{\text{basel}}, \ell_t^{\text{basel}})\mathrm{d}t\right] = \mathbb{E}_{\tilde{T}}\left[\int_{t=\tilde{T}}^{\infty} \mathrm{e}^{-\hat{\rho}(t-\tilde{T})}u(c_t^{\text{pol}}, \ell_t^{\text{pol}})\mathrm{d}t\right]$ holds for $\tilde{T} = [1960, 1980] + 20$.

under the policies compared to the baseline. The skill type s is on the x-axes and $\phi(s)$ on the y-axes, where $\phi(s)$ satisfies

$$\mathbb{E}_{\tilde{T}}\left[\int_{t=\tilde{T}}^{\infty} \mathrm{e}^{-\hat{\rho}(t-\tilde{T})}u\left(\phi(s)\cdot c_t^{\text{basel}}, \ell_t^{\text{basel}}\right)\mathrm{d}t\right] = \mathbb{E}_{\tilde{T}}\left[\int_{t=\tilde{T}}^{\infty} \mathrm{e}^{-\hat{\rho}(t-\tilde{T})}u\left(c_t^{\text{pol}}, \ell_t^{\text{pol}}\right)\mathrm{d}t\right]$$

for cohorts that are of skill type s, and that join the labor market in $\tilde{T} = 1980$ and 2000 and are therefore born around 1960 and 1980, respectively.[32]

While the heterogeneity across birth cohorts is of minor importance, the heterogeneity across skill types is quite significant. On the one hand, low- and medium-skilled workers gain from the policies. This is both because they directly benefit from the programs by collecting the loans (wage replacements) in case they switch to the abstract occupation, and because of the advantageous general equilibrium effects on routine and manual wages discussed above. Quantitatively, medium-skilled types born in 1980 can increase expected lifetime consumption by about 0.3% (1.0%) under the optimal loan (wage replacement) program. High-skilled workers, on the other hand, lose in terms of welfare. They usually work in the abstract occupation their entire life and hence never directly benefit from the policies. However, they have to pay higher labor taxes (as everyone else in the economy) and they are harmed by the lower abstract wage.

[32]Given that the evolution of wages in the model lags behind that in the data by 10 to 15 years (Section 4.6) the exact timing of the welfare gains pertaining to the different cohorts should be interpreted with some care. In particular, given that abstract wages started to rise earlier in the data than in the model, the welfare gains of the policies might be concentrated among earlier cohorts than the analysis here suggests.

Output, ICT capital, and labor productivity

That the policies lead to lower wages of high-skilled and higher wages of low- and middle-skilled workers could imply that the redistributive character of the general equilibrium effects is the only source of aggregate welfare gains. As Figure 1.12 shows, however, the policies also increase output and therefore go beyond redistribution. The left panel plots relative deviations of output Y_t under the policies from its values in the baseline transition. During the first years, especially under the wage replacement program, effects on output are slightly negative. This owes to higher labor taxes and the additional switching of experienced workers which destroys human capital in the short term. After 2000, both policies have positive effects on output which persist far into the future. Averaged over the years 1980 to 2050, both policies cause output to rise by about 0.05% compared to the baseline, with a peak of about 0.1% during the later years of the transition.[33] This is in line with the results from the extended simple model discussed at the end of Section 2. Alleviating the borrowing constraints leads to a more efficient allocation of labor across occupations and therefore benefits output.

The middle panel shows that the policies lead to a crowding in of ICT capital of up to 1% (3%). As there is less supply of routine work, which is a substitute to ICT capital, and more supply of abstract work, which is a complement, firms optimally react by using more ICT capital in production when the policies are in place. Finally, as the right panel shows, the policies have a positive impact on labor productivity, as measured by output divided by hours worked. This is driven by the increase in ICT capital as well as the improved allocation of workers across occupations, i.e. more abstract and less routine labor. These two factors are intimately linked, given the complementarities inherent to the aggregate production function (1.2): the reallocation of labor towards the abstract occupation only leads to productivity gains if accompanied by an increase in ICT capital.[34]

When introducing the policies in the steady state of the model, output does not go up (see Figure 1.27 in the appendix). It stays largely unchanged under the loan program and even falls when the wage replacement payments are introduced. While ICT capital is again crowded in as more routine workers switch into the abstract occupation, the policies do not lead to an increase in labor productivity. The fact that output does not rise when the policies are introduced in steady state indicates that the detrimental effects of the policies discussed above outweigh any positive effect from an improved allocation of workers across occupations and the increase in capital.

[33] The dip in 2025 is explained by the fact that the programs are phased out in that year (see design of the policies). This causes some last workers to take advantage of the policies, leading to a spike in labor taxes to finance this (see Figure 1.26 in the appendix).

[34] When fixing the capital stock at its values under the baseline transition and only letting labor inputs $N_{j,t}$ adjust to their new values under the policies, the effect on output is negative. Only in conjunction with the increase in ICT capital do the policies cause output gains.

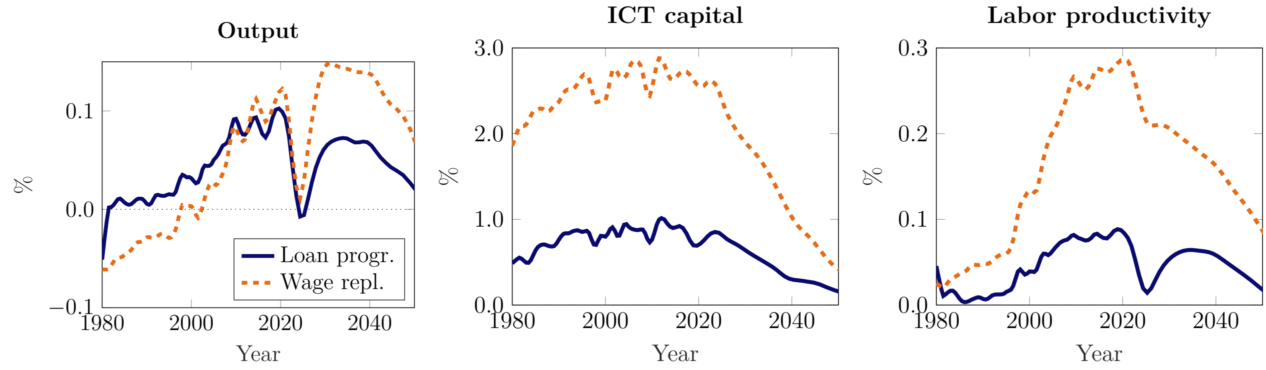


Figure 1.12: Deviations of variables under optimal policies relative to baseline transition.
Notes: Left: Output Y_t. Middle: ICT capital $K_{ict,t}$. Right: Labor prod. $Y_t / \left(\int_i \ell_t^* \, \mathrm{d}\Gamma_t(i)\right)$.

7 Conclusion

The US labor market has undergone a profound structural change over the recent decades. Caused by a falling price of ICT capital, jobs intensive in routine tasks have become automated and wages for routine labor have suffered. As a result of this, a significant fraction of the US workforce has reallocated into occupations that perform work complementary to ICT capital.

I demonstrated that borrowing constraints are important to understand the output and welfare consequences of this labor market adjustment. In a simplified model, I highlighted that occupational choices are distorted towards staying in today's attractive occupations when households are constrained in borrowing against future income compared to the case when they are not. Empirically, I documented three facts. First, while routine workers who switched to the manual or abstract occupations saw faster wage growth on average than those who stayed in the routine occupations, a large share of them faced initial wage losses. Second, at least a third of routine workers is hand-to-mouth, i.e. possesses very few liquid assets to smooth consumption. Third, not being hand-to-mouth has been predictive of leaving routine occupations.

Building on these findings, I developed a general equilibrium model with incomplete markets and occupation-specific human capital. Along a transition during which relative wages in the routine occupation decline, households close to the borrowing constraint are more reluctant to leave the routine occupation and start over new in the manual or abstract occupation than households with many liquid assets. This causes employment shares to be subdued in the baseline economy compared to a counterfactual scenario in which all households behave like the liquidity-rich.

I used the model to study two policies: a loan and a transfer program. Both programs alleviate the borrowing constraint of experienced routine workers who leave their former occupation. These workers lose occupation-specific human capital and hence often incur wage losses upon switching while enjoying wage gains only in the long run. When targeted towards workers who switch to the abstract occupation, both policies have positive effects on social welfare. This owes in large part to general equilibrium effects, as wages of low-income households rise and additional ICT capital and an improved allocation of labor raise output. My results indicate that handing out transfers is preferable to loans, since the former better targets those routine workers who have a long-run interest in staying in the abstract occupation.

I made a number of simplifications to keep the analysis tractable. All agents in the model have perfect foresight over future technological growth and therefore wage paths in all occupations. Assuming some degree of uncertainty over future wages or even myopia when it comes to occupational choices might be more realistic. Also, it could be interesting to study the incentives of firms more closely. Some firms might retrain routine workers for new tasks instead of firing them once they are replaced by machines (Dauth et al., 2021). This seems to be especially relevant for understanding the transition process in European countries, though not as much in the US where firing costs are low (Mukoyama et al., 2021). I captured moves up (and down) the occupational ladder in a reduced-form, exogenous way by adding a shock (η) to the productivity process of workers. In general, I abstracted from explicitly modeling labor market frictions to focus on the role of borrowing constraints in the labor market transition. All these aspects are left for future work.

In the future, more jobs will likely become automated that so far have been performed by machines (Edin et al., 2021). Moreover, Albanesi and Kim (2021) and Chernoff and Warman (2020) argue that the Covid-19 pandemic might accelerate the trend in automation. Machines are not susceptible to falling ill due to a virus, and safeguarding the workplace against the spread of diseases further adds to the cost of employing labor relative to machines. These considerations could spur investment in labor-saving technology. Also, while in the past a separation of occupations into routine and non-routine has been important to understand employment growth, the pandemic could make on-the-job contact to other human beings as well as flexibility to work remotely relevant task dimensions. Kaplan, Moll, et al. (2020) document that workers in social-intensive, low-flexibility occupations tend to hold few liquid assets. This suggests that the same mechanism I highlighted in this paper will slow down the reallocation of labor into less contact-intensive and more flexible occupations.

Appendices

A Proofs for Simple Model

Note that the model of Section 2 can be more generally described as a model where the household arrives in T either with some prior occupation $j_{T-1} \in \{r, a\}$ or as a newborn with no prior occupation, $j_{T-1} = \emptyset$. The prior occupation is then the only relevant state variable of the household. If she works in the same occupation for two consecutive periods, she is an experienced household ($\bar{h}$). If she switches the occupation or if she has no prior occupation (e.g. because she newly arrived to the labor market) she is inexperienced ($\underline{h}$). Therefore, $h_t = \bar{h} \cdot \mathbf{1}_{\{j_t = j_{t-1}\}} + \underline{h} \cdot \mathbf{1}_{\{j_t \neq j_{t-1}\}}$.

Experienced routine worker Let me first state the proposition regarding the occupational choice of a worker who was priorly employed in the routine occupation, i.e. with $j_{T-1} = r$. The potential occupational choices are:

1. $\{j_T = r, j_{T+1} = r\} \rightarrow \{y_T = \bar{h}, y_{T+1} = \bar{h}\}$
2. $\{j_T = r, j_{T+1} = a\} \rightarrow \{y_T = \bar{h}, y_{T+1} = \omega \cdot \underline{h}\}$
3. $\{j_T = a, j_{T+1} = a\} \rightarrow \{y_T = \underline{h}/\omega, y_{T+1} = \omega \cdot \bar{h}\}$
4. $\{j_T = a, j_{T+1} = r\} \rightarrow \{y_T = \underline{h}/\omega, y_{T+1} = \underline{h}\}$

Proposition 2. *If not borrowing-constrained, the household switches to the abstract occupation in $t = T$ iff*

$$1 + r \leq \omega - \frac{\max\left\{\omega, \bar{h}/\underline{h}\right\} - 1}{\bar{h}/\underline{h} - 1/\omega}$$

In this case, she is a net borrower in $t = T$.
If borrowing-constrained, the household never switches to the abstract occupation in $t = T$.

The result resembles the one for the inexperienced worker in Section 2. This was to be expected: the only difference in the case of the experienced routine worker is that choosing the abstract occupation in period T is even more costly because of the already accumulated occupation-specific human capital. This also explains why a growing human capital spread $\bar{h}/\underline{h}$ does not unambiguously make the condition in Proposition 2 more likely to hold (like it did in Proposition 1), as it further raises the cost of leaving the routine occupation. As was the case for inexperienced workers, experienced workers who are borrowing-constrained never choose the abstract occupation in T.

A.1 Proofs of Propositions 1 and 2

Unconstrained households I first prove the result regarding the unconstrained household of Proposition 1 in the main text.

For the household to choose the abstract occupation in period T, it must hold that option 3 yields higher lifetime income than the maximum of options 1 and 2, i.e.

$$\underline{h} + \frac{1}{1+r}\max\{\bar{h}, \omega\underline{h}\} \leq \underline{h}/\omega + \frac{\omega}{1+r}\bar{h}$$

Case 1: $\bar{h}/\underline{h} \geq \omega$

$$\begin{aligned}
\underline{h} + \frac{1}{1+r}\bar{h} &\leq \underline{h}/\omega + \frac{\omega}{1+r}\bar{h} \\
(1+r)\left(1 - \frac{1}{\omega}\right) &\leq (\omega - 1)\bar{h}/\underline{h} \\
1 + r &\leq \frac{\omega - 1}{1 - \frac{1}{\omega}}\bar{h}/\underline{h} = \omega\bar{h}/\underline{h} = \omega(\bar{h}/\underline{h} - 1)\frac{\bar{h}/\underline{h}}{\bar{h}/\underline{h} - 1}
\end{aligned}$$

Case 2: $\bar{h}/\underline{h} < \omega$

$$\begin{aligned}
\underline{h} + \frac{\omega}{1+r}\underline{h} &\leq \underline{h}/\omega + \frac{\omega}{1+r}\bar{h} \\
(1+r)\left(1 - \frac{1}{\omega}\right) &\leq \omega(\bar{h}/\underline{h} - 1) \\
1 + r &\leq \omega\frac{\bar{h}/\underline{h} - 1}{1 - \frac{1}{\omega}} = \omega(\bar{h}/\underline{h} - 1)\frac{\omega}{\omega - 1}
\end{aligned}$$

Rearranging yields the desired result, which I reprint here for convenience:

$$1 + r \leq \omega(\bar{h}/\underline{h} - 1)\min\left\{\frac{\bar{h}/\underline{h}}{\bar{h}/\underline{h} - 1}, \frac{\omega}{\omega - 1}\right\}.$$

The proof of Proposition 2 regarding the unconstrained households follows analogously.

It is also straightforward to verify that whenever the unconstrained household chooses option 3, she borrows against future income. Too see this, note that consumption in period T is $c_T = \frac{1}{1+\beta}\left(\underline{h}/\omega + \frac{\omega}{1+r}\bar{h}\right)$, and hence savings in period T are $s = \frac{1}{1+\beta}\left(\beta\underline{h}/\omega - \frac{\omega}{1+r}\bar{h}\right)$. Assume for a contradiction that savings in T are positive

Case 1: $\bar{h}/\underline{h} \geq \omega$

$$\begin{aligned}
& \beta\underline{h}/\omega > \frac{\omega}{1+r}\bar{h} \\
\Rightarrow \quad & \underline{h}/\omega > \frac{\omega}{1+r}\bar{h} \\
\Rightarrow \quad & 1+r > \omega^2\bar{h}/\underline{h} > \omega\bar{h}/\underline{h} \quad \lightning
\end{aligned}$$

Case 2: $\bar{h}/\underline{h} < \omega$

$$\begin{aligned}
& \beta\underline{h}/\omega > \frac{\omega}{1+r}\bar{h} \\
\Rightarrow \quad & \underline{h}/\omega > \frac{\omega}{1+r}\bar{h} \\
\Rightarrow \quad & 1+r > \omega^2\bar{h}/\underline{h} > \omega^2\frac{\bar{h}/\underline{h}-1}{\omega-1} = \omega\frac{\bar{h}/\underline{h}-1}{1-\frac{1}{\omega}} \quad \lightning
\end{aligned}$$

Both contradictions arise from the conditions derived in the first part of the proof.

Constrained households Next I prove that if the household is prevented from borrowing, she never chooses option 3.

Suppose first that in the optimal solution the borrowing constraint is not binding. In this case, the problem of the constrained household is identical to that of the unconstrained household, whose solution I have just derived. We know, however, that if parameters are such that option 3 (choosing abstract in $t = T$) is optimal she is a net borrower, a contradiction to the assumption that the constraint is not binding.

Now suppose the borrowing constraint is binding and the constrained household consumes her current income in both periods. For her to find it optimal to choose option 3 it must hold that

$$\ln(\underline{h}) + \beta\max\left\{\ln\left(\bar{h}\right), \ln(\omega\underline{h})\right\} \leq \ln(\underline{h}/\omega) + \beta\ln\left(\omega\bar{h}\right)$$

Case 1: $\bar{h}/\underline{h} \geq \omega$

$$\begin{aligned}
\ln(\underline{h}) + \beta\ln\left(\bar{h}\right) &\leq \ln(\underline{h}) - \ln(\omega) + \beta\ln(\omega) + \beta\ln\left(\bar{h}\right) \\
0 &\leq \beta - 1 \quad \lightning
\end{aligned}$$

The contradiction follows from the fact that I have assumed (strict) discounting, i.e. $\beta < 1$.
Case 2: $\bar{h}/\underline{h} < \omega$

$$\begin{aligned}
\ln(\underline{h}) + \beta\ln(\omega) + \beta\ln(\underline{h}) &\leq \ln(\underline{h}) - \ln(\omega) + \beta\ln(\omega) + \beta\ln\left(\bar{h}\right) \\
\ln(\omega) &\leq \beta\ln\left(\bar{h}/\underline{h}\right) \\
\omega &\leq (\bar{h}/\underline{h})^{\beta} \quad \lightning
\end{aligned}$$

The proof of Proposition 2 regarding the constrained households follows analogously.

A.2 Time-varying aggregate productivity

Consider now the introduction of an additional variable $Z > 1$ that augments all wages in period $T + 1$, i.e.

$$\begin{aligned} w_{r,T} &= 1, & w_{r,T+1} &= Z \\ w_{a,T} &= \frac{1}{\omega}, & w_{a,T+1} &= Z\omega \end{aligned}$$

The following proposition holds regarding the occupational choice of an inexperienced household, and can be proved analogously to Proposition 1 above.

Proposition 3. *If not borrowing-constrained, the household chooses the abstract occupation in $t = T$ iff*

$$\frac{1+r}{Z} \leq \omega(\bar{h}/\underline{h} - 1) \min\left\{\frac{\bar{h}/\underline{h}}{\bar{h}/\underline{h} - 1}, \frac{\omega}{\omega - 1}\right\}.$$

In this case, she is a net borrower in $t = T$.
If borrowing-constrained, the household never chooses the abstract occupation in $t = T$.

This result is identical to Proposition 1, except for the fact that the gross interest rate $(1+r)$ is dampened by the factor Z. Put differently, a higher Z makes choosing the abstract occupation today more likely. Intuitively, if wages today are relatively low compared to tomorrow, the opportunity cost of choosing the abstract occupation today is also relatively low. This is in line with empirical evidence on the business cycle patters of the aggregate decline in routine labor (Hershbein and Kahn, 2018; Jaimovich and Siu, 2020).[35]

B Simple Model with Endogenous Wages and Skill Distribution

In the following I extend the simple model of Section 2 in two ways. First, I endogenize wages. Second, I allow for a continuum of households with heterogeneous skill type. The extended environment allows me to characterize the efficient allocation of labor across occupations when there is relative wage growth in the abstract relative to the routine occupation. I then study the allocation in the decentralized equilibrium, once when households can freely borrow against future income and once when everyone is hand-to-mouth. In the former case, the allocation of labor coincides with that chosen by the planner. In the latter, relative to the first-best, too few households work in the abstract and too many in the routine occupation.

[35] The cyclical patterns of the hand-to-mouth share have so far, to the best of my knowledge, not been studied in detail. Note, however, that the share of hand-to-mouth households in the US economy depicted in Figure 1.16 or in Kaplan, Violante, et al. (2014) does not display any notable pro- or countercyclical pattern.

B.1 Environment

Time t is discrete and runs until period $T+1$.[36] There exists a representative firm that uses two inputs, routine labor N_r and abstract labor N_a. It produces output according to the production function $F(N_r, N_a; q)$ where q is the exogenous level of technology.

There exists a continuum of measure one of households. Households derive utility from consumption $u(c)$ and provide one unit of labor inelastically in one of the two occupations. A household is characterized by her fixed skill type s and her previous period's occupation j_{t-1}. Households face a constant probability of dying $\pi \in (0,1)$. A dead household is replaced by a newborn who has no previous occupation, $j_{t-1} = \emptyset$. All households die with certainty after period $T+1$. Conditional on survival, households discount future utility by a factor $\beta \in (0, \frac{1}{1-\pi})$. Denote the effective discount rate of households $\hat{\beta} \equiv \beta \cdot (1-\pi)$, and hence $\hat{\beta} \in (0,1)$.

Skill is distributed in the population according to some probability distribution function $s \sim g(s)$. Households accumulate occupation-specific human capital $h \in \{\underline{h}, \bar{h}\}$, with $0 < \underline{h} < \bar{h}$. If a household stays in the same occupation in two consecutive periods, i.e. $j_t = j_{t-1}$ she is an experienced worker in period t ($h_t = \bar{h}$). If she switches occupations or if she is a newborn with no prior occupation, she is inexperienced ($h_t = \underline{h}$), as $j_t \neq j_{t-1}$. The functions $\phi_j(s)$ with $j \in \{r, a\}$ map skill into productivity in the respective occupation. Hence if a household works in occupation j at time t, her labor supply is $\phi_j(s) \cdot h_t$. I make the following assumptions:

Assumptions 1.

- *$F(\cdot)$ is continuously differentiable, with $F_{N_j} > 0$ for $j = r, a$*
- *Conditional on factor inputs N_j, q raises the marginal product of abstract labor relative to that of routine labor, $\frac{\partial \frac{F_{N_a}}{F_{N_r}}}{\partial q} > 0$*
- *$u(\cdot)$ is twice continuously differentiable, with $u_c > 0$, $u_{cc} < 0$*
- *$g(s)$ has positive mass on the whole real line: $\forall s \in \mathbb{R}, g(s) > 0$*
- *$\phi(s) \equiv \frac{\phi_a(s)}{\phi_r(s)}$ is continuously differentiable, with $\phi'(s) > 0$*

The last assumption implies that high-skilled types have a comparative advantage in the abstract occupation, low-skilled types in the routine occupation.[37]

[36] The framework could be extended to an infinite time horizon. Ending in $T+1$, however, makes the problem below a more tractable two-period problem (T and $T+1$), which greatly simplifies the exposition and algebra.

[37] This assumption is common in the literature (Cortes, 2016; vom Lehn, 2020) and it is met in the full model of the main text, where I calibrate $\phi_j(s) = \exp(a_j s)$ with $0 = a_m < a_r < a_a$.

B.2 Social planner

I begin by studying the first-best allocation of labor that arises as the solution to a social planner's problem. I have abstracted from physical capital as it is not crucial for the analysis. Instead, I assume that the economy can transfer resources on aggregate to the next period by lending and borrowing from the rest of the world at rate r, with $1+r \leq 1/\beta$. Denote aggregate holdings of the foreign asset at time t by B_t. All debt has to be repaid at the end of period $T+1$.

No technological change

Assume first that there is no technological change: $\forall t, q_t = q$. In period T, the planner needs to allocate consumption and an occupation to each household for the remaining two periods T and $T+1$. Note first that since I have assumed concave utility, the planner assigns the same level of consumption C_t to each household. Second, the assumption on the productivity functions $\phi_j(s)$ and the fact that marginal products of labor are constant imply that there is no occupational mobility and that there exists a cut-off S that separates the skill space into low-skilled (who work in the routine occupation their entire life) and high-skilled (abstract occupation).

$$
\begin{aligned}
\max_{S,C_T,C_{T+1}} \; & u(C_T) + \beta \cdot u(C_{T+1}) \\
& \text{subject to:} \\
& C_T + \frac{C_{T+1}}{1+r} = Y_T + \frac{Y_{T+1}}{1+r} + B_T \\
& \text{and for } t = T, T+1: \\
& Y_t = Y = F(N_r, N_a; q) \\
& N_r = \int_{-\infty}^{S} \left[\pi \cdot (\phi_r(s) \cdot \underline{h}) + (1-\pi) \cdot (\phi_r(s) \cdot \bar{h})\right] g(s)\mathrm{d}s \\
& N_a = \int_{S}^{\infty} \left[\pi \cdot (\phi_a(s) \cdot \underline{h}) + (1-\pi) \cdot (\phi_a(s) \cdot \bar{h})\right] g(s)\mathrm{d}s
\end{aligned}
$$

The optimal skill cut-off, which maximizes output, is characterized by:

$$
\begin{aligned}
& F_{N_r} \cdot \phi_r(S) = F_{N_a} \cdot \phi_a(S) \\
\Leftrightarrow\; & F_{N_r} = F_{N_a} \cdot \phi(S) \qquad (1.8)
\end{aligned}
$$

Intuitively, this condition ensures that the worker with skill type $s = S$ is equally productive in both occupations.

Technological change

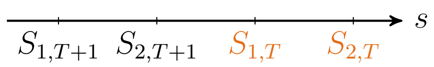
Now assume that technology is constant up until time T. In period T, however, it becomes known that $q_{T+1} > q_T$ and hence the marginal product of abstract labor relative to that of routine labor grows. It is thus optimal to reallocate workers to the abstract occupation, i.e. even those with skill below the optimal skill cut-off absent technological change S. Moreover, it might be optimal to start doing so already in period T such that workers can start building up occupation-specific human capital.

The planner now needs to choose four cut-off levels, one for each combination of human capital and time period. For $t = T, T+1$, she needs to decide up to which skill level $S_{1,t}$ inexperienced workers work in the routine occupation. She also needs to decide up to which level $S_{2,t}$ experienced routine workers keep working in the routine occupation. Note that in any optimal allocation it must hold that $S_{1,t} \leq S_{2,t}$. In words, young workers are at least as likely to enter the abstract occupation as experienced routine workers, since the latter have already accumulated occupation-specific human capital.[38]

$$
\begin{aligned}
\max_{\{S_{1,t},S_{2,t},C_t\}_{t=T,T+1}} \quad & u(C_T) + \beta \cdot u(C_{T+1}) \\
& \text{subject to:} \\
& C_T + \frac{C_{T+1}}{1+r} = Y_T + \frac{Y_{T+1}}{1+r} \\
& Y_t = F(N_{r,t}, N_{a,t}; q_t), \text{ for } t = T, T+1 \\
& N_{r,T} = \pi \int_{-\infty}^{S_{1,T}} (\phi_r(s)\underline{h}) g(s) \mathrm{d}s + (1-\pi) \int_{-\infty}^{S_{2,T}} (\phi_r(s)\bar{h}) g(s) \mathrm{d}s \\
& N_{a,T} = \pi \int_{S_{1,T}}^{\infty} (\phi_a(s)\underline{h}) g(s) \mathrm{d}s + (1-\pi) \left[\int_{S}^{\infty} (\phi_a(s)\bar{h}) g(s) \mathrm{d}s + \int_{S_{2,T}}^{S} (\phi_a(s)\underline{h}) g(s) \mathrm{d}s \right]
\end{aligned}
$$

and subject to labor market clearing in $T+1$, for which one needs to differentiate three cases, depending on the ordering of $S_{1,T}$, $S_{2,T}$ and $S_{2,T+1}$:[39]

<u>Case 1:</u>



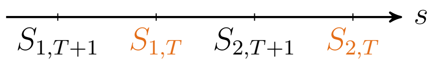
[38] This is easy to prove by contradiction. Assume that $S_{1,t} > S_{2,t}$. But then one can move the cut-offs closer together, achieving the same amount of abstract labor supply but more routine labor supply. This will become evident formally when I derive the optimal cut-offs below.

[39] The position of $S_{1,T+1}$ relative to $S_{1,T}$ and $S_{2,T}$ does not affect the labor market clearing conditions.

$$N_{r,T+1} = \pi \int_{-\infty}^{S_{1,T+1}} (\phi_r(s)\underline{h})g(s)\mathrm{d}s + (1-\pi) \int_{-\infty}^{S_{2,T+1}} (\phi_r(s)\bar{h})g(s)\mathrm{d}s$$

$$\begin{aligned} N_{a,T+1} = \pi \int_{S_{1,T+1}}^{\infty} (\phi_a(s)\underline{h})g(s)\mathrm{d}s + (1-\pi)\Bigg[\pi \int_{S_{2,T+1}}^{S_{1,T}} (\phi_a(s)\underline{h})g(s)\mathrm{d}s + (1-\pi) \int_{S_{2,T+1}}^{S_{2,T}} (\phi_a(s)\underline{h})g(s)\mathrm{d}s \\ + (1-\pi) \int_{S_{2,T}}^{\infty} (\phi_a(s)\bar{h})g(s)\mathrm{d}s + \pi \int_{S_{1,T}}^{\infty} (\phi_a(s)\bar{h})g(s)\mathrm{d}s\Bigg] \end{aligned}$$

<u>Case 2:</u>



$$N_{r,T+1} = \pi \int_{-\infty}^{S_{1,T+1}} (\phi_r(s)\underline{h})g(s)\mathrm{d}s + (1-\pi)\left[(1-\pi) \int_{-\infty}^{S_{2,T+1}} (\phi_r(s)\bar{h})g(s)\mathrm{d}s + \pi \int_{-\infty}^{S_{1,T}} (\phi_r(s)\bar{h})g(s)\mathrm{d}s\right]$$

$$\begin{aligned} N_{a,T+1} = \pi \int_{S_{1,T+1}}^{\infty} (\phi_a(s)\underline{h})g(s)\mathrm{d}s + (1-\pi)\Bigg[(1-\pi) \int_{S_{2,T+1}}^{S_{2,T}} (\phi_a(s)\underline{h})g(s)\mathrm{d}s \\ + (1-\pi) \int_{S_{2,T}}^{\infty} (\phi_a(s)\bar{h})g(s)\mathrm{d}s + \pi \int_{S_{1,T}}^{\infty} (\phi_a(s)\bar{h})g(s)\mathrm{d}s\Bigg] \end{aligned}$$

<u>Case 3:</u>

$$N_{r,T+1} = \pi \int_{-\infty}^{S_{1,T+1}} (\phi_r(s)\underline{h})g(s)\mathrm{d}s + (1-\pi)\left[(1-\pi) \int_{-\infty}^{S_{2,T}} (\phi_r(s)\bar{h})g(s)\mathrm{d}s + \pi \int_{-\infty}^{S_{1,T}} (\phi_r(s)\bar{h})g(s)\mathrm{d}s\right]$$

$$N_{a,T+1} = \pi \int_{S_{1,T+1}}^{\infty} (\phi_a(s)\underline{h})g(s)\mathrm{d}s + (1-\pi)\left[(1-\pi) \int_{S_{2,T}}^{\infty} (\phi_a(s)\bar{h})g(s)\mathrm{d}s + \pi \int_{S_{1,T}}^{\infty} (\phi_a(s)\bar{h})g(s)\mathrm{d}s\right]$$

For Case 1, this yields the following optimality conditions that implicitly characterize the cut-off levels:

$$
\begin{aligned}
&[S_{1,T}] : \frac{(1-\pi)\cdot F_{N_{a,T+1}}\cdot \phi(S_{1,T})\cdot(\bar{h}/\underline{h}-1)}{F_{N_{r,T}} - F_{N_{a,T}}\cdot\phi(S_{1,T})} = 1+r \\
&[S_{2,T}] : \frac{(1-\pi)\cdot F_{N_{a,T+1}}\cdot \phi(\tilde{S}_{2,T})\cdot(\bar{h}/\underline{h}-1)}{F_{N_{r,T}}\cdot(\bar{h}/\underline{h}) - F_{N_{a,T}}\cdot\phi(\tilde{S}_{2,T})} = 1+r, \quad S_{2,T} = \min\{\tilde{S}_{2,T}, S\} \\
&[S_{1,T+1}] : F_{N_{r,T+1}} = F_{N_{a,T+1}}\cdot\phi(S_{1,T+1}) \\
&[S_{2,T+1}] : F_{N_{r,T+1}}\cdot(\bar{h}/\underline{h}) = F_{N_{a,T+1}}\cdot\phi(\tilde{S}_{2,T+1}), \quad S_{2,T+1} = \min\{\tilde{S}_{2,T+1}, S\}
\end{aligned}
$$

Focus on the first condition, i.e. for $S_{1,T}$. Intuitively, choosing the optimal cut-offs in period T is an investment decision. Putting an additional household to work in the abstract occupation reduces output today, a cost which appears in the denominator (remember that the denominator would be zero for $\phi(S)$ and becomes positive for $S_{1,T} < S$, see equation (1.8)). This cost must equal the benefit of having an additional abstract worker in period $T+1$, which appears in the numerator.

It is easy to see that the skill cut-off for inexperienced workers in period T is smaller than it was absent technological change, i.e. $S_{1,T} < S$, since $\phi(s)$ is continuous and monotone in s and

$$
\begin{aligned}
\lim_{S_{1,T}\to S^-} \frac{(1-\pi)\cdot F_{N_{a,T+1}}\cdot \phi(S_{1,T})\cdot(\bar{h}/\underline{h}-1)}{F_{N_{r,T}} - F_{N_{a,T}}\cdot\phi(S_{1,T})} &= +\infty \\
\lim_{S_{1,T}\to -\infty} \frac{(1-\pi)\cdot F_{N_{a,T+1}}\cdot \phi(S_{1,T})\cdot(\bar{h}/\underline{h}-1)}{F_{N_{r,T}} - F_{N_{a,T}}\cdot\phi(S_{1,T})} &= 0\,.
\end{aligned}
$$

In Case 2, $S_{1,T}$ is instead characterized by

$$
[S_{1,T}] : \frac{(1-\pi)\cdot(F_{N_{a,T+1}}\cdot \phi(S_{1,T}) - F_{N_{r,T+1}})\cdot(\bar{h}/\underline{h})}{F_{N_{r,T}} - F_{N_{a,T}}\cdot\phi(S_{1,T})} = 1+r
$$

while the remaining cut-offs are as in Case 1. In Case 3, $S_{2,T}$ is instead characterized by

$$
[S_{2,T}] : \frac{(1-\pi)\cdot(F_{N_{a,T+1}}\cdot \phi(\tilde{S}_{2,T}) - F_{N_{r,T+1}})\cdot(\bar{h}/\underline{h})}{F_{N_{r,T}}\cdot(\bar{h}/\underline{h}) - F_{N_{a,T}}\cdot\phi(\tilde{S}_{2,T})} = 1+r, \quad S_{2,T} = \min\{\tilde{S}_{2,T}, S\}
$$

while the remaining cut-offs are as in Case 2.

B.3 Competitive equilibrium

Next I show that the same cut-off skill levels that I have just derived as the optimal solution to the planner's problem arise from a decentralized equilibrium in which households can freely borrow against future income and in which the representative firm maximizes profits. Suppose first that $\forall t, q_t = q$ such that both wages are constant over time. In this case, it is optimal for households to choose the occupation in which they are most productive and work there their entire life. Hence, absent wage growth the cut-off S from the planner's problem again separates the skill space into routine and abstract workers and the allocation of labor is efficient.

Let me therefore turn to the more interesting case of solving the problem with technological change, $q_{T+1} > q_T$. Let b_t denote the amount of assets a household currently holds. Denoting by $w_{j,t}$ the wage a household earns when working in occupation j, a household of type (s, j_{T-1}, b_T) solves:[40]

$$\max_{\{j_t, c_t\}_{t=T,T+1}} \quad u(c_T) + \hat{\beta} \cdot u(c_{T+1})$$

$$\text{subject to:}$$

$$c_T + \frac{c_{T+1}}{(1+r)/(1-\pi)} = y(s, j_{T-1}, j_T) + \frac{y(s, j_T, j_{T+1})}{(1+r)/(1-\pi)} + b_T$$

Focus first on an inexperienced household, $j_{T-1} = \emptyset$. Her lifetime income is

$$y_T + \frac{y_{T+1}}{(1+r)/(1-\pi)} = \begin{cases} \underline{h} \cdot \phi_r(s) \cdot w_{r,T} + \frac{\bar{h} \cdot \phi_r(s) \cdot w_{r,T+1}}{(1+r)/(1-\pi)} & \text{if } j_T = r, j_{T+1} = r \\ \underline{h} \cdot \phi_r(s) \cdot w_{r,T} + \frac{\underline{h} \cdot \phi_a(s) \cdot w_{a,T+1}}{(1+r)/(1-\pi)} & \text{if } j_T = r, j_{T+1} = a \\ \underline{h} \cdot \phi_a(s) \cdot w_{a,T} + \frac{\bar{h} \cdot \phi_a(s) \cdot w_{a,T+1}}{(1+r)(1-\pi)} & \text{if } j_T = a, j_{T+1} = a \end{cases} \tag{1.9}$$

I want to determine the skill type $S^*_{1,T}$ that is exactly indifferent between working in the routine and the abstract occupation in period T.

<u>Case A:</u> (corresponds to Case 1 of the planner)

$$\bar{h} \cdot \phi_r(S^*_{1,T}) \cdot w_{r,T+1} \leq \underline{h} \cdot \phi_a(S^*_{1,T}) \cdot w_{a,T+1}$$

In this case, the skill type $S^*_{1,T}$ is determined by the following equation:

$$\underline{h} \cdot \phi_r(S^*_{1,T}) \cdot w_{r,T} + \frac{\underline{h} \cdot \phi_a(S^*_{1,T}) \cdot w_{a,T+1}}{(1+r)/(1-\pi)} = \underline{h} \cdot \phi_a(S^*_{1,T}) \cdot w_{a,T} + \frac{\bar{h} \cdot \phi_a(S^*_{1,T}) \cdot w_{a,T+1}}{(1+r)/(1-\pi)}$$

$$\Rightarrow \quad \frac{(1-\pi) \cdot w_{a,T+1} \cdot \phi(S^*_{1,T}) \cdot (\bar{h}/\underline{h} - 1)}{w_{r,T} - \phi(S^*_{1,T}) \cdot w_{a,T}} = 1 + r$$

<u>Case B:</u> (corresponds to Cases 2 and 3 of the planner)

$$\bar{h} \cdot \phi_r(S^*_{1,T}) \cdot w_{r,T+1} > \underline{h} \cdot \phi_a(S^*_{1,T}) \cdot w_{a,T+1}$$

[40] The assumption here is that households may die with positive or negative asset holdings. The foreign creditor realizes this and augments the gross interest rate households have to pay by a factor $\frac{1}{1-\pi}$. This way, aggregate debt can be honored in the last period since $(1+r)\int s_T = (1-\pi)\int (c_{T+1} - w_{T+1})$, where $s_T = w_T - c_T$.

In this case, the skill type $S^*_{1,T}$ is determined by:

$$\underline{h} \cdot \phi_r(S^*_{1,T}) \cdot w_{r,T} + \frac{\bar{h} \cdot \phi_r(S^*_{1,T}) \cdot w_{r,T+1}}{(1+r)/(1-\pi)} = \underline{h} \cdot \phi_a(S^*_{1,T}) \cdot w_{a,T} + \frac{\bar{h} \cdot \phi_a(S^*_{1,T}) \cdot w_{a,T+1}}{(1+r)/(1-\pi)}$$

$$\Rightarrow \quad \frac{(1-\pi) \cdot (w_{a,T+1} \cdot \phi(S^*_{1,T}) - w_{r,T+1}) \cdot (\bar{h}/\underline{h})}{w_{r,T} - \phi(S^*_{1,T}) \cdot w_{a,T}} = 1 + r$$

Since the optimal behavior of the representative firm ensures that wages equal the marginal product of labor, $w_{j,t} = F_{N_{j,t}}$, the condition that characterizes $S^*_{1,T}$ coincides with that for $S_{1,T}$ derived in the planner's problem, so the two cut-offs are equal. The same property can be shown in an analogous fashion for the cut-off $S_{1,T+1}$ as well as for the cut-offs $S_{2,T}$ and $S_{2,T+1}$ when analyzing the problem of experienced routine workers with $j_{T-1} = r$. Hence, the allocation of labor in the decentralized economy is efficient.

For future reference, let me define wage growth in occupation j as

$$\omega_j \equiv \frac{w_{j,T+1}}{w_{j,T}}$$

and let me explicitly solve for $\phi(S_{1,T})$:

<u>Case A:</u>

$$\phi(S_{1,T}) = \frac{w_{r,T}}{w_{a,T}} \cdot \frac{1}{1 + \frac{1-\pi}{1+r} \cdot \omega_a \cdot (\bar{h}/\underline{h} - 1)}$$

<u>Case B:</u>

$$\phi(S_{1,T}) = \frac{w_{r,T}}{w_{a,T}} \cdot \frac{1 + \frac{1-\pi}{1+r} \cdot (\bar{h}/\underline{h}) \cdot \omega_r}{1 + \frac{1-\pi}{1+r} \cdot (\bar{h}/\underline{h}) \cdot \omega_a}$$

B.4 Competitive equilibrium with hand-to-mouth households

I now analyze the equilibrium allocation when all households in the economy are hand-to-mouth, i.e. they consume their current income in both periods. Consuming current income would endogenously arise as the optimal choice of all households if wages grew over time in both occupations and households held zero assets and were exogenously prevented from borrowing against future income.[41]

Suppose first that $\forall t, q_t = q$ such that both wages are constant over time. In this case, it is optimal for hand-to-mouth households to choose the occupation in which they are most productive and work there their entire life. Hence, absent wage growth the cut-off S from the planner's problem again separates the skill space into routine and abstract workers and the allocation of labor is efficient, as in the case without borrowing constraints. I therefore turn to analyzing the problem with technological change, i.e. $q_{T+1} > q_T$. The problem of a

[41] Under these assumptions, the proof of this is analogous to the one provided in Appendix A.1 for the simpler model.

household with type (s, j_{T-1}) is:

$$\max_{j_T, j_{T+1}} u(y_T) + \hat{\beta} \cdot u(y_{T+1})$$

Note that occupational choices lead to the same implications for income y_t as shown in (1.9). To make progress, I make the following assumptions in what follows:

Assumptions 2.

- $u(c) = \ln(c)$
- $\omega_a \geq 1$
- $\omega_a > \omega_r$

The last two assumptions are fairly weak. $\omega_a \geq 1$ implies that abstract wages grow (weakly). This provides an incentive to borrow against future income which households in this subsection cannot do, leading to an inefficiency.[42] $\omega_a > \omega_r$ indicates that wage growth is stronger in the abstract than in the routine occupation. Note that this requirement is a bit stronger than the second bullet point of Assumptions 1, which only implied that *conditional on fixed factor inputs* the marginal product of abstract labor grows in comparison to that of routine labor. Essentially, $\omega_a > \omega_r$ requires that the reduction of labor supply in the routine occupation (increase of labor supply in the abstract occupation) does not overturn the growth in relative abstract wages that results from technological change.

As before, focus on the decision of an inexperienced household, i.e. with $j_{T-1} = \emptyset$.
Case A:

$$\bar{h} \cdot \phi_r(S^+_{1,T}) \cdot w_{r,T+1} \leq \underline{h} \cdot \phi_a(S^+_{1,T}) \cdot w_{a,T+1}$$

In this case, the skill type $S^+_{1,T}$ that is indifferent between working in the routine and the abstract occupation in period T is determined by:

$$\begin{aligned}
\ln\Big(\phi_r(S^+_{1,T}) \cdot w_{r,T} \cdot \underline{h}\Big) + \hat{\beta} \cdot \ln\Big(\phi_a(S^+_{1,T}) \cdot w_{a,T+1} \cdot \underline{h}\Big) \\
= \ln\Big(\phi_a(S^+_{1,T}) \cdot w_{a,T} \cdot \underline{h}\Big) \\
+ \hat{\beta} \cdot \ln\Big(\phi_a(S^+_{1,T}) \cdot w_{a,T+1} \cdot \bar{h}\Big) \\
(\phi_r(S^+_{1,T}) \cdot w_{r,T} \cdot \underline{h}) \cdot (\phi_a(S^+_{1,T}) \cdot w_{a,T+1} \cdot \underline{h})^{\hat{\beta}} &= (\phi_a(S^+_{1,T}) \cdot w_{a,T} \cdot \underline{h}) \cdot (\phi_a(S^+_{1,T}) \cdot w_{a,T+1} \cdot \bar{h})^{\hat{\beta}} \\
\phi(S^+_{1,T}) &= \frac{w_{r,T}}{w_{a,T}} \cdot (\bar{h}/\underline{h})^{-\hat{\beta}}
\end{aligned}$$

Case B:

$$\bar{h} \cdot \phi_r(S^+_{1,T}) \cdot w_{r,T+1} > \underline{h} \cdot \phi_a(S^+_{1,T}) \cdot w_{a,T+1}$$

[42] In fact, all results that follow hold also under the weaker condition $\omega_a \geq \beta \cdot (1 + r)$.

In this case, $S^{+}_{1,T}$ is determined by:

$$(\phi_r(S^{+}_{1,T}) \cdot w_{r,T} \cdot \underline{h}) \cdot (\phi_r(S^{+}_{1,T}) \cdot w_{r,T+1} \cdot \bar{h})^{\hat{\beta}} = (\phi_a(S^{+}_{1,T}) \cdot w_{a,T} \cdot \underline{h}) \cdot (\phi_a(S^{+}_{1,T}) \cdot w_{a,T+1} \cdot \bar{h})^{\hat{\beta}}$$

$$\phi(S^{+}_{1,T}) = \frac{w_{r,T}}{w_{a,T}} \cdot \left(\frac{\omega_r}{\omega_a}\right)^{\frac{\hat{\beta}}{1+\hat{\beta}}}$$

To show that fewer workers choose to work in the abstract occupation in T when borrowing constraints bind than when borrowing is unrestricted I next prove that $\phi(S^{+}_{1,T}) > \phi(S^{*}_{1,T})$, which by monotonicity of $\phi(s)$ implies $S^{+}_{1,T} > S^{*}_{1,T}$.

<u>Case A:</u>

$$\phi(S^{*}_{1,T}) < \phi(S^{+}_{1,T})$$

$$\frac{1}{1 + \frac{1-\pi}{1+r} \cdot \omega_a \cdot (\bar{h}/\underline{h} - 1)} < (\bar{h}/\underline{h})^{-\hat{\beta}}$$

$$0 < \underbrace{1 + \frac{1-\pi}{1+r} \cdot \omega_a \cdot (\bar{h}/\underline{h} - 1) - (\bar{h}/\underline{h})^{\hat{\beta}}}_{\Lambda(\bar{h}/\underline{h})}$$

For this inequality to be fulfilled it needs to hold that $\forall(\bar{h}/\underline{h}) > 1, \Lambda(\bar{h}/\underline{h}) > 0$. Note that $\Lambda(1) = 0$. Hence it suffices to show that $\forall(\bar{h}/\underline{h}) \geq 1, \Lambda'(\bar{h}/\underline{h}) > 0$.

$$\Lambda'(\bar{h}/\underline{h}) = \frac{1-\pi}{1+r} \cdot \omega_a - \hat{\beta}(\bar{h}/\underline{h})^{\hat{\beta}-1} > 0$$

$$\underbrace{\omega_a}_{\geq 1} > \underbrace{\beta \cdot (1+r)}_{\leq 1} \cdot \underbrace{(\bar{h}/\underline{h})^{\hat{\beta}-1}}_{<1}$$

<u>Case B:</u>

$$\phi(S^{*}_{1,T}) < \phi(S^{+}_{1,T})$$

$$\frac{1 + \frac{1-\pi}{1+r} \cdot (\bar{h}/\underline{h}) \cdot \omega_r}{1 + \frac{1-\pi}{1+r} \cdot (\bar{h}/\underline{h}) \cdot \omega_a} < \left(\frac{\omega_r}{\omega_a}\right)^{\frac{\hat{\beta}}{1+\hat{\beta}}}$$

Define $x \equiv \frac{\omega_a}{\omega_r}$ and take the log on both sides.

$$0 < \underbrace{\ln\Big(1 + \frac{1-\pi}{1+r} \cdot (\bar{h}/\underline{h}) \cdot \omega_r \cdot x\Big) - \ln\Big(1 + \frac{1-\pi}{1+r} \cdot (\bar{h}/\underline{h}) \cdot \omega_r\Big) - \frac{\hat{\beta}}{1+\hat{\beta}} \ln(x)}_{H(x)}$$

For this inequality to be fulfilled it needs to hold that $\forall x > 1, H(x) > 0$. Note that $H(1) = 0$.

Hence it suffices to show that $\forall x \geq 1, H'(x) > 0$.

$$
\begin{aligned}
H'(x) = \frac{\frac{1-\pi}{1+r} \cdot (\bar{h}/\underline{h}) \cdot \omega_r}{1 + \frac{1-\pi}{1+r} \cdot (\bar{h}/\underline{h}) \cdot \omega_r \cdot x} - \frac{\hat{\beta}}{1+\hat{\beta}} \frac{1}{x} &> 0 \\
\frac{\frac{1-\pi}{1+r} \cdot (\bar{h}/\underline{h}) \cdot \omega_r}{\frac{1}{x} + \frac{1-\pi}{1+r} \cdot (\bar{h}/\underline{h}) \cdot \omega_r} &> \frac{\hat{\beta}}{1+\hat{\beta}} \\
\frac{1+\hat{\beta}}{\hat{\beta}} &> 1 + \frac{1}{x \cdot (\bar{h}/\underline{h}) \cdot \frac{1-\pi}{1+r} \cdot \omega_r} \\
1 &> \frac{\hat{\beta}}{x \cdot (\bar{h}/\underline{h}) \cdot \frac{1-\pi}{1+r} \cdot \omega_r} \\
\underbrace{\omega_a \cdot (\bar{h}/\underline{h})}_{>1} &> \underbrace{\beta \cdot (1+r)}_{\leq 1}
\end{aligned}
$$

This completes the proof that $S^{+}_{1,T} > S^{*}_{1,T} = S_{1,T}$ and a similar result can be shown to hold for the cut-off $S_{2,T}$ when analyzing the problem of an experienced routine worker. Hence, fewer skill types work in the abstract occupation in T when households are hand-to-mouth than is optimal. This in turn leads to a shortage of experienced abstract workers in period $T+1$.

C Data Sources and Measurement

C.1 Broad occupational groups

Based on earlier work by Dorn (2009), Autor and Dorn (2013) map Census Occupation Codes from 1950 to 2005 into a time-consistent set of occupations. They then form six broad occupational groups, based on the task content in each occupation which they in turn derive from the US Department of Labor's Document of Occupational Titles. Throughout this paper, I use the following categorization of these six groups into manual, routine and abstract occupations, which coincides with the concepts used in Jaimovich and Siu (2020) and in the appendix of vom Lehn (2020):

- Abstract:
  - management/professional/technical/financial sales/public security occupations
- Routine:
  - administrative support and retail sales occupations
  - precision production and craft occupations
  - machine operators, assemblers and inspectors

    - transportation/construction/mechanics/mining/agricultural occupations

- Manual:
    - low-skill services

C.2 PSID

Sample selection

I use waves 1976 to 2017 of the PSID and I follow Cortes (2016) in terms of sample selection and when defining variables. In particular, Cortes (2016) uses only the core PSID sample, drops all individuals who have never been a household head or a wife and only keeps observations aged 16–64. Note that while Cortes (2016) uses data only up until 2007, I include data until 2017, i.e. five more waves of the PSID. Whenever using real wages (or changes thereof) in the analysis, I follow Cortes (2016) in excluding observations with log hourly real wages below 1.10$ or above 54.60$ in terms of 1979 dollars (3.87$ and 192.30$ in terms of 2019 dollars). This is relevant for the results in Section 3.2 and when estimating returns to tenure in the calibration. Like Cortes (2016), I use the Consumer Price Index for All Urban Consumers: All Items in U.S. City Average (CPIAUCSL, downloaded from the Federal Reserve Economic Database) to deflate nominal quantities. The empirical analysis in Section 3.2 further only uses male observations to make the results comparable to Cortes (2016)'s.

As mentioned in the main text, in order to be consistent throughout this paper I marginally deviate from Cortes (2016) when assigning the broad occupation to individuals. While Cortes (2016) groups occupations according to the classification given in Acemoglu and Autor (2011), I form groups from the six broad groups listed in Autor and Dorn (2013) as described above. The two concepts yield very similar results, and Cortes (2016) shows in his appendix that using a classification that is closer to the one used by Autor and Dorn (2013) leads to very similar results as his baseline.

Wage changes of switchers vs stayers

To obtain estimates of how wages of switchers from routine to abstract and manual occupations evolved relative to routine stayers at horizon h, Cortes (2016) performs the following regression

$$\Delta_h \ln(wage_{it}) = \beta_m^h \cdot D_{imt} + \beta_a^h \cdot D_{iat} + \gamma_t^h + u_{iht}$$

where $\Delta_h \ln(wage_{it})$ is the change in the log real hourly wage of individual i between years t and $t+h$, D_{ijt} is a dummy variable equal to zero if the individual was working in the routine occupation in both t and $t+1$ (or t and $t+2$ for $h \geq 2$) and one if she switched from the routine occupation to occupation j, γ_t captures time fixed effects, and u_{iht} is a mean zero error term. The solid blue lines in Figure 1.13 visualize his estimates of β_m^h (left panel) and β_a^h (right panel). The dotted blue lines show my own estimates, which employ the slightly

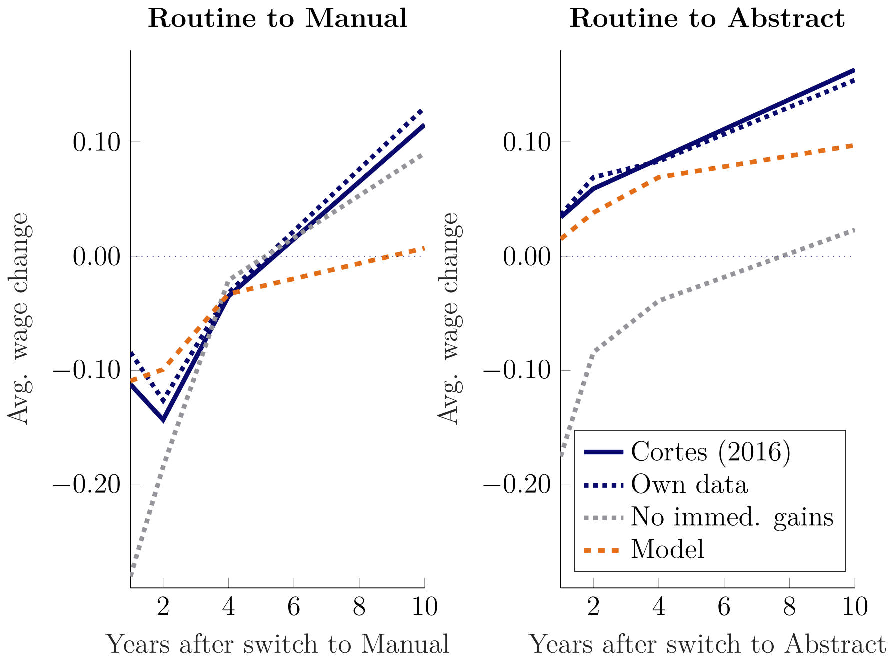


Figure 1.13: Average log hourly real wages of switchers from routine to manual (left) and to abstract (right) compared to routine stayers.

Notes: Blue solid lines are the estimates from Cortes (2016, Table 3). "Own data" replicates his analysis. "No immed. gains" excludes all individuals from the sample who saw positive wage changes from leaving the routine occupation on impact. For obtaining the model values I perform the same regression as Cortes (2016), only using data from a synthetic panel of 10,000 households, simulated between 1980 and 2020 from the full model of Section 4.

different occupational groups as well as the additional five waves of the PSID. As can be seen, the estimates are nearly identical to Cortes (2016)'s.

The grey dotted lines show the estimates when deleting all observations from the sample who see immediate wage gains (i.e. between t to $t+1$) when leaving the routine occupation. Since the PSID went to bi-annual frequency after 1997, so that I do not observe year-on-year wage changes in the later waves, I restrict the sample to observations prior to 1997 for computing these estimates. Estimating the wage effects only on workers who see initial wage losses when leaving the routine occupations obviously lowers the average wage gains obtained from switching, but, as can be seen, even this subset of workers saw faster wage growth than the stayers. The dashed red lines show the estimates from the synthetic panel simulated from the full model of Section 4. While the initial average wage change (i.e. after year 1) is targeted in the calibration, the remaining horizons are untargeted, and still provide a relatively good fit to Cortes (2016)'s estimates.

Returns to tenure

As specified in the main text, when estimating returns to tenure in the calibration, I follow the procedure outlined in Cortes (2016, Section VI C), using years 1981 to 2017 only (see Figure 1.14). I only deviate from him by leaving females in the sample (as employment shares used in the calibration also relate to both males and females (Autor and Dorn, 2013)), and

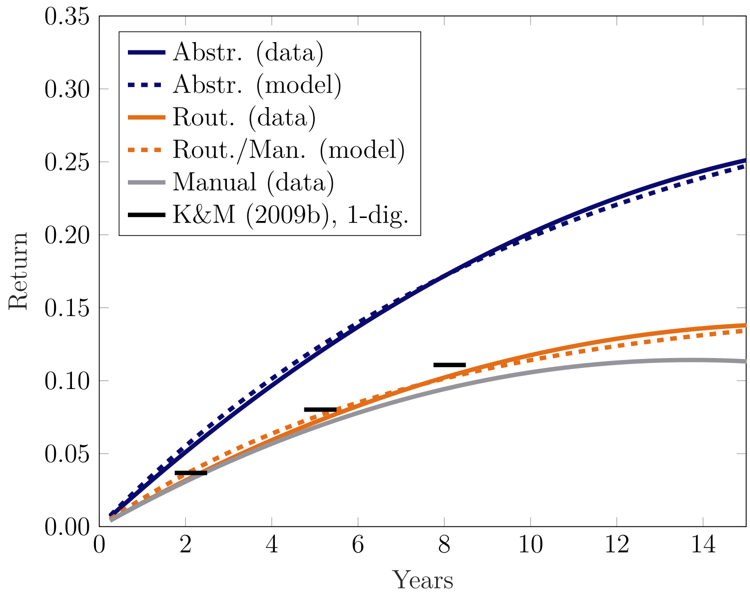


Figure 1.14: Returns to occupational tenure, estimated from PSID data (1981-2017).
Notes: I estimate $y_{it} = \sum_j D_{ijt}\left(\beta_{j1}Ten_{ijt} + \beta_{j2}Ten^2_{ijt} + \gamma_{ij}\right) + \delta X_{it} + u_{it}$, where y_{it} is the log real wage of individual i at time t, D_{ijt} is a dummy indicating whether the individual was working in occupation $j \in \{m, r, a\}$ at time t, Ten_{ijt} is occupational tenure, γ_{ij} is an occupation fixed effect for each individual, and X_{it} are controls (unionization and marital status, region of residence, year dummies and year-occupation dummies). I then calibrate the parameters of the Poisson process in the model $(\bar{h}^j - \underline{h})$ and λ^j_h to obtain the closest fit (in a mean squared sense) to the data over the first 20 years.

by only assigning a value for broad occupational tenure to individuals once I have observed them making an occupational switch. Following Kambourov and Manovskii (2009b), every time a person indicates working in a different occupational group than in her most previous report, Cortes (2016) considers the person to have made a switch and occupational tenure is reset. For individuals who are observed for the first time in the survey, Cortes (2016) constructs estimates of occupational tenure by setting it equal to the person's stated tenure at her employer or in her current position. This is of course an imperfect estimate of tenure in the broad occupation and adds significant measurement error to the variable. As I have many additional observations in comparison to Cortes (2016) because of the five more waves of the PSID that I use, I choose to limit myself to individuals whom I have observed to make a switch and for whom I can therefore construct the measure relatively cleanly.

Robustness checks for Section 3.2

Tables 1.5 and 1.6 below separately use switching from routine to abstract and from routine to manual as the dependent (dummy) variable. They show that the main result from Table 1.1, i.e. that being hand-to-mouth is negatively associated with a future switch, is largely driven by switches to the abstract occupation. In fact, the point estimates for switches to manual are all positive, however not statistically significant different from zero.

Table 1.7 differs from the baseline in that I add females to the sample. In this case, the point estimates of β become statistically indistinguishable from zero. Estimating the models

Table 1.5: Switching decision, switch from routine to abstract occupations.

	(1)	(2)	(3)	(4)
HtM	-3.96***	-2.81**	-0.18***	-0.13***
	(1.10)	(1.09)	(0.048)	(0.049)
Occ. tenure	-1.24***	-1.15***	-0.085***	-0.080***
	(0.090)	(0.086)	(0.0080)	(0.0078)
Skill		11.4***		0.49***
		(1.12)		(0.049)
Controls	Yes	Yes	Yes	Yes
Observations	4904	4904	4904	4904
Model	OLS	OLS	Probit	Probit

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Data are from the PSID, years 1999–2017. Sample selection is as in Cortes (2016). Definition of HtM status is as in Kaplan, Violante, et al. (2014). Dependent variable is whether or not individual leaves routine for the abstract occ. between t and $t + 2$ (coded as 0 or 100). Skill is a dummy for whether the individual has received more than 12 years of education. Occ. tenure is years of uninterrupted tenure in broad occupation. Additional controls are: region dummies, age, age squared, unionization status, married status, year. Standard errors are clustered at the individual level.

Table 1.6: Switching decision, switch from routine to manual occupations.

	(1)	(2)	(3)	(4)
HtM	0.92	0.67	0.11	0.074
	(0.56)	(0.56)	(0.070)	(0.071)
Occ. tenure	-0.19***	-0.21***	-0.040***	-0.044***
	(0.043)	(0.044)	(0.010)	(0.011)
Skill		-2.42***		-0.33***
		(0.57)		(0.074)
Controls	Yes	Yes	Yes	Yes
Observations	4904	4904	4904	4904
Model	OLS	OLS	Probit	Probit

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Data are from the PSID, years 1999–2017. Sample selection is as in Cortes (2016). Definition of HtM status is as in Kaplan, Violante, et al. (2014). Dependent variable is whether or not individual leaves routine for the manual occ. between t and $t + 2$ (coded as 0 or 100). Skill is a dummy for whether the individual has received more than 12 years of education. Occ. tenure is years of uninterrupted tenure in broad occupation. Additional controls are: region dummies, age, age squared, unionization status, married status, year. Standard errors are clustered at the individual level.

Table 1.7: Switching decision, including females.

	(1)	(2)	(3)	(4)
HtM	-0.59	-0.065	-0.018	0.00055
	(0.86)	(0.86)	(0.030)	(0.031)
Occ. tenure	-1.46***	-1.40***	-0.077***	-0.075***
	(0.070)	(0.069)	(0.0046)	(0.0046)
Skill		6.85***		0.23***
		(0.85)		(0.030)
Controls	Yes	Yes	Yes	Yes
Observations	10048	10048	10048	10048
Model	OLS	OLS	Probit	Probit

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Data are from the PSID, years 1999–2017. Sample selection is as in Cortes (2016), but includes females. Definition of HtM status is as in Kaplan, Violante, et al. (2014). Dependent variable is whether or not individual leaves manual occ. between t and $t + 2$ (coded as 0 or 100). Skill is a dummy for whether the individual has received more than 12 years of education. Occ. tenure is years of uninterrupted tenure in broad occupation. Additional controls are: region dummies, age, age squared, unionization status, married status, year. Standard errors are clustered at the individual level.

only with women in the sample (estimates not shown) reveals that this result is driven by women being more likely to switch to the manual (low-skilled services) occupations when the household has few liquid assets. The reason for this could be intra-household insurance mechanisms that I do not model. In particular, it might be easier for women to flexibly adjust working hours (e.g. from part- to full-time work) if they work in the services occupations (e.g. bar tending, child care). This could explain their switching behavior when the household finds itself in a financially dire situation.

In Table 1.8 I split up the independent variable “HtM” into an interaction term between being hand-to-mouth and experiencing wage gains upon switching. Consistent with the intuition from the quantitative model it is especially those switches that entail wage losses (second row) that hand-to-mouth agents are especially likely to avoid.

Lastly, I ask whether being hand-to-mouth also predicts switches out of abstract (Table 1.9) and out of manual (Table 1.10) occupations. None of the estimated coefficients on the hand-to-mouth dummy are statistically different from zero. This is in line with the quantitative model of Section 4, in which liquid assets are an especially strong determinant of switching behavior only among routine workers.

Table 1.8: Switching decision, estimated with interaction terms.

	(1)	(2)	(3)	(4)
HtM=0 × wage_gain=1	2.47	2.61*	0.098	0.11*
	(1.56)	(1.56)	(0.061)	(0.062)
HtM=1 × wage_gain=0	-4.00**	-3.17	-0.17*	-0.14
	(1.99)	(2.00)	(0.087)	(0.088)
HtM=1 × wage_gain=1	-0.25	0.83	-0.015	0.028
	(1.67)	(1.67)	(0.066)	(0.067)
Occ. tenure	-1.43***	-1.36***	-0.083***	-0.079***
	(0.098)	(0.095)	(0.0072)	(0.0071)
Skill		9.01***		0.33***
		(1.21)		(0.045)
Controls	Yes	Yes	Yes	Yes
Observations	4904	4904	4904	4904
Model	OLS	OLS	Probit	Probit

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Data are from the PSID, years 1999–2017. Sample selection is as in Cortes (2016). Definition of HtM status is as in Kaplan, Violante, et al. (2014). Dependent variable is whether or not individual leaves routine for the manual occ. between t and $t + 2$ (coded as 0 or 100). Skill is a dummy for whether the individual has received more than 12 years of education. Occ. tenure is years of uninterrupted tenure in broad occupation. Additional controls are: region dummies, age, age squared, unionization status, married status, year. Standard errors are clustered at the individual level.

Table 1.9: Switching decision, exit from abstract occupations.

	(1)	(2)	(3)	(4)
HtM	0.21	-0.18	0.0023	-0.013
	(1.17)	(1.17)	(0.045)	(0.045)
Occ. tenure	-1.08***	-0.98***	-0.053***	-0.049***
	(0.072)	(0.072)	(0.0044)	(0.0044)
Skill		-10.6***		-0.35***
		(1.51)		(0.048)
Controls	Yes	Yes	Yes	Yes
Observations	5567	5567	5567	5567
Model	OLS	OLS	Probit	Probit

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Data are from the PSID, years 1999–2017. Sample selection is as in Cortes (2016). Definition of HtM status is as in Kaplan, Violante, et al. (2014). Dependent variable is whether or not individual leaves abstract occ. between t and $t + 2$ (coded as 0 or 100). Skill is a dummy for whether the individual has received more than 12 years of education. Occ. tenure is years of uninterrupted tenure in broad occupation. Additional controls are: region dummies, age, age squared, unionization status, married status, year. Standard errors are clustered at the individual level.

Table 1.10: Switching decision, exit from manual occupations.

	(1)	(2)	(3)	(4)
HtM	1.17	1.11	0.025	0.024
	(2.96)	(2.98)	(0.085)	(0.085)
Occ. tenure	-1.84***	-1.85***	-0.063***	-0.063***
	(0.31)	(0.32)	(0.013)	(0.013)
Skill		-0.44		-0.012
		(2.92)		(0.083)
Controls	Yes	Yes	Yes	Yes
Observations	1136	1136	1136	1136
Model	OLS	OLS	Probit	Probit

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Data are from the PSID, years 1999–2017. Sample selection is as in Cortes (2016). Definition of HtM status is as in Kaplan, Violante, et al. (2014). Dependent variable is whether or not individual leaves manual occ. between t and $t + 2$ (coded as 0 or 100). Skill is a dummy for whether the individual has received more than 12 years of education. Occ. tenure is years of uninterrupted tenure in broad occupation. Additional controls are: region dummies, age, age squared, unionization status, married status, year. Standard errors are clustered at the individual level.

Definition of being hand-to-mouth

In classifying households as either being hand-to-mouth or not I follow Kaplan, Violante, et al. (2014). For a detailed record of how they define income, liquid and illiquid wealth, see pages 122-3 in Kaplan, Violante, et al. (2014). I use information on income and wealth on determining whether a household is hand-to-mouth as they do and I detail this procedure in the next subsection.

C.3 SCF

I use waves 1989 to 2019 of the SCF and I follow Kaplan, Violante, et al. (2014) in terms of sample selection and when defining variables. In particular, I consider all households whose head is aged 22–79, and discard those who report negative labor income, and those whose only positive income stems from self-employment. I refer the reader to Kaplan, Violante, et al. (2014) for further details.

Definition of being hand-to-mouth

I define the liquid asset holdings of a household m_{it} as the sum of cash, money market accounts, checking/savings/call accounts, prepaid cards, directly held stocks, bonds, non-money-market mutual funds, minus revolving consumer debt. Income y_{it} collects labor earnings, regular private transfers (e.g. child support, alimony), and public transfers (e.g. unemployment benefits, food stamps, Social Security Income). Income y_{it} corresponds to bi-weekly income, as this is the most common frequency of payment in the US (Kaplan, Violante, et al., 2014).

Households are considered hand-to-mouth if and only if one of the following two conditions is true:

$$0 \leq m_{it} \leq \frac{y_{it}}{2}$$

or

$$m_{it} \leq \frac{y_{it}}{2} - \underline{m}_{it} \ ,$$

where $\underline{m}_{it}$ corresponds to a household's borrowing constraint, which, in line with Kaplan, Violante, et al. (2014)'s baseline definition, I assume to be one times monthly income.

Mapping of occupations

From 1989 until 2001, the SCF used the 3-digit 1980 and 1990 Census occupation codes, which were both very similar. The public-use files contain identifiers for whether the household head worked in one of six broad occupational groups. I map these six groups as closely as possible to the three groups used throughout this paper (manual, routine, abstract) using the consistent occupation-classification scheme proposed in the data appendix of Dorn (2009) (*Occ1990dd*).

Table 1.11: Mapping of SCF occupation dummies into broad occupational groups.

Broad Group	Manual	Routine	Abstract
SCF occupation dummies	3	2, 4, 5, 6	1

Table 1.11 lists the assignment of SCF groups to the three broad occupational groups. While the overlap is very large and the assignment therefore unambiguous for the SCF groups 1 and 3 to 6, group 2 contains some abstract and some routine occupations (codes 203–389 of the 1980 Census classification). I therefore further look at the employment shares in each of these occupations provided by Autor and Dorn (2013) for the years 1990 and 2000. A significantly smaller fraction of workers in this group was employed in abstract occupations (8-9%, codes 203–258) than in routine occupations (21–22%, codes 274–389). Hence, I classify all workers of group 2 as routine, which also explains why I slightly overestimate the routine employment share in Figure 1.15 in the first years.

From 2004 onward, the SCF used the 2000 and 2010 Census codes, which were very similar to each other but quite distinct from the earlier Census codes. The SCF then assigned households again into six broad groups. While the overlap between the three broad occupational groups and the six SCF groups is somewhat weaker than prior to 2004, I still find that the assignment displayed in Table 1.11 yields the closest mapping.

To demonstrate that my classification of occupations is close to that of Autor and Dorn (2013), Figure 1.15 plots the self-constructed shares of employment obtained from the SCF next to those in Autor and Dorn (2013). While I am overestimating (underestimating) slightly the share of workers employed in routine (abstract) occupations in the early 1990s, the time series are relatively closely aligned towards the early 2000s.

Time series of hand-to-mouth shares

Figure 1.16 plots the shares of households that I classify as being hand-to-mouth separately for each broad occupational group as well as unconditionally for all households. Two results stand out. First, there is a clear ordering of hand-to-mouth shares by occupational group, with workers in the manual occupations most likely to be hand-to-mouth, and abstract workers least likely. This is perhaps not surprising, given that abstract workers are usually the ones earning the highest incomes, manual workers earning the lowest incomes, and routine workers in between.

Second, routine workers, while not as likely as manual workers to be hand-to-mouth, are still more likely to be so than the average US household. Across all years, the probability of routine workers to be hand-to-mouth was on average 35%, higher than the average probability across all households of 29%. At each single point in time, the two series differ by three to seven percentage points.

Figures 1.17 and 1.18 further document that the high share of hand-to-mouth households among the routine workers is driven to a large extent by the high prevalence of wealthy

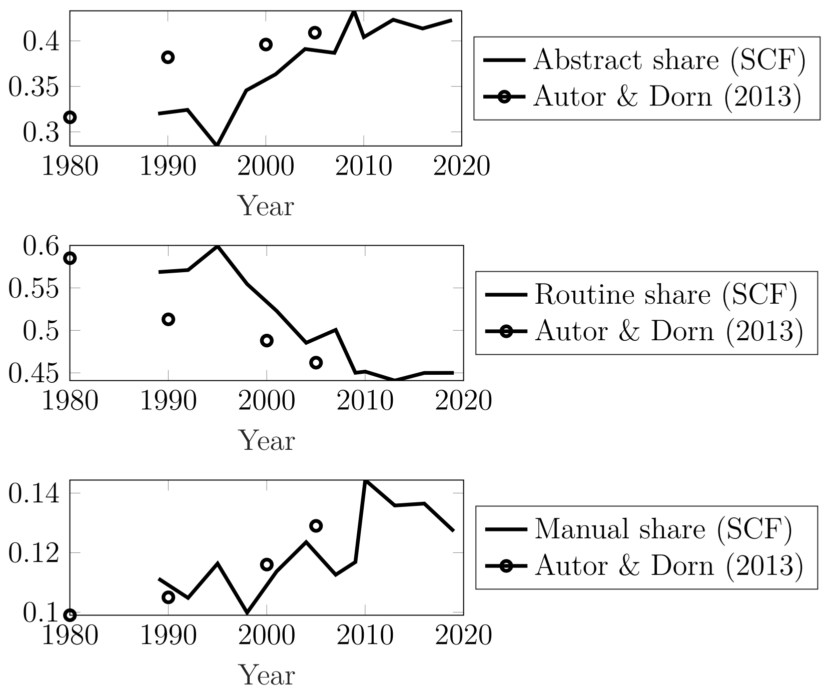


Figure 1.15: Employment shares in the SCF and in Autor and Dorn (2013).

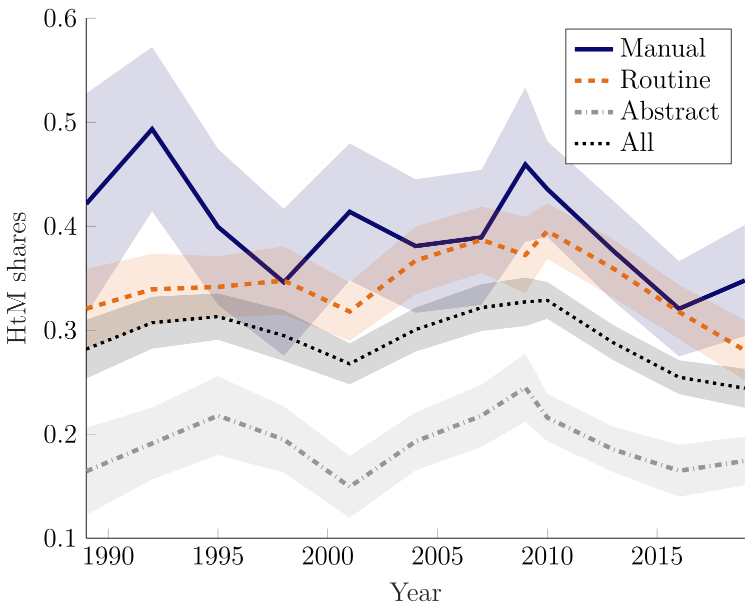


Figure 1.16: Hand-to-mouth shares in the three broad occupational groups.
Notes: Confidence intervals are at the 95% level. "All" refers to households whose head works in either of the three occupations, i.e. is employed.

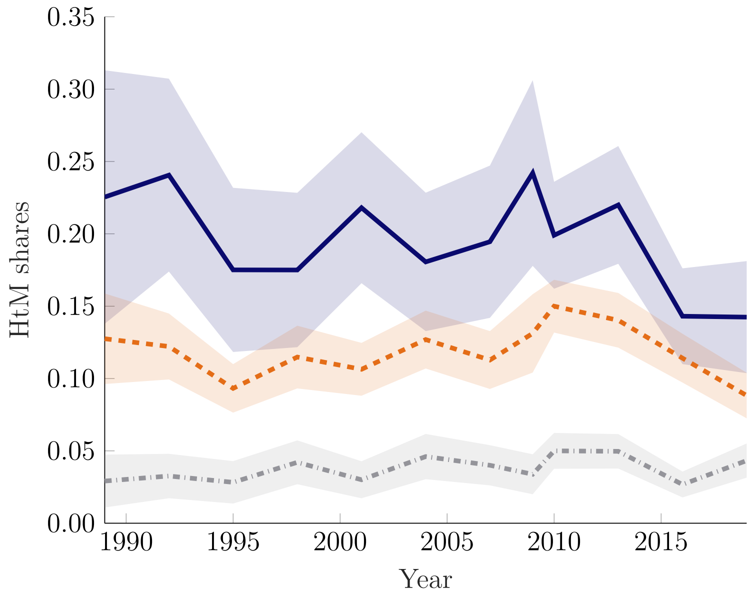


Figure 1.17: Poor hand-to-mouth shares in the three broad occupational groups.
Notes: Confidence intervals are at the 95% level.

hand-to-mouth households among them. These are households who are by definition hand-to-mouth, but own positive illiquid wealth such as equity in houses or indirect stock holdings (Kaplan, Violante, et al., 2014). In all years between 1995 and 2016, routine workers were more likely to be wealthy hand-to-mouth than either abstract or manual workers.

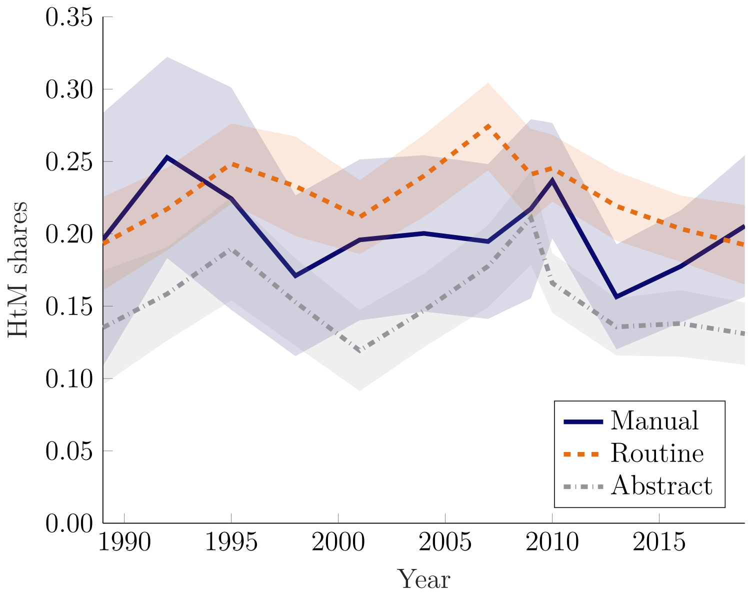


Figure 1.18: Wealthy hand-to-mouth shares in the three broad occupational groups.
Notes: Confidence intervals are at the 95% level.

D Details of the Full Model

D.1 FOCs of the representative firm

$$w_m = K_s^{\alpha}(1-\alpha)\tilde{Y}^{-\alpha}\mu_m\left(\frac{\tilde{Y}}{N_m}\right)^{\frac{1}{\gamma_m}} \tag{1.10}$$

$$w_r = \Omega(1-\mu_a)R^{\frac{\gamma_a-\gamma_r}{\gamma_a\gamma_r}}\mu_r N_r^{\frac{-1}{\gamma_r}} \tag{1.11}$$

$$w_a = \Omega\mu_a N_a^{\frac{-1}{\gamma_a}} \tag{1.12}$$

$$r_{ict} = \Omega(1-\mu_a)R^{\frac{\gamma_a-\gamma_r}{\gamma_a\gamma_r}}(1-\mu_r)K_{ict}^{\frac{-1}{\gamma_r}} \tag{1.13}$$

$$r_s = \alpha K_s^{\alpha-1}\tilde{Y}^{1-\alpha} \tag{1.14}$$

where

$$\Omega = K_s^{\alpha}(1-\alpha)\tilde{Y}^{\frac{-\alpha}{\gamma_m}}(1-\mu_m)\left(\mu_a N_a^{\frac{\gamma_a-1}{\gamma_a}} + (1-\mu_a)R^{\frac{\gamma_a-1}{\gamma_a}}\right)^{\frac{\gamma_m-\gamma_a}{(\gamma_a-1)\gamma_m}}$$

$$\tilde{Y} = \left[\mu_m N_m^{\frac{\gamma_m-1}{\gamma_m}} + (1-\mu_m)\left[\mu_a N_a^{\frac{\gamma_a-1}{\gamma_a}} + (1-\mu_a)R^{\frac{\gamma_a-1}{\gamma_a}}\right]^{\frac{\gamma_a(\gamma_m-1)}{(\gamma_a-1)\gamma_m}}\right]^{\frac{\gamma_m}{\gamma_m-1}}$$

D.2 Hamilton-Jacobi-Bellman equation

The solution to the problem of a household with state η_k is characterized by

$$\begin{aligned}
\hat{\rho} V_t(s,\eta_k,\epsilon,h,j,m,\tilde{k}) = &\max_{c,\ell,d} u(c,\ell) \\
&+V_{m,t}(s,\eta_k,\epsilon,h,j,m,\tilde{k})\left[(1-\tau)w_j\ell y + \mathbf{1}_{m<0}\kappa m + T - d - \chi(d,\tilde{k}) - c\right] \\
&+V_{\tilde{k},t}(s,\eta_k,\epsilon,h,j,m,\tilde{k})(r_k\tilde{k}+d) \\
&+\lambda_\eta\left[V_t(s,\eta_{-k},\epsilon,h,j,m,\tilde{k}) - V_t(s,\eta_k,\epsilon,h,j,m,\tilde{k})\right] \\
&+V_{\epsilon,t}(s,\eta_k,\epsilon,h,j,m,\tilde{k})(-\beta_\epsilon\epsilon) \\
&+\lambda_\epsilon\int_{-\infty}^{\infty}\left[V_t(s,\eta_k,x,h,j,m,\tilde{k}) - V_t(s,\eta_k,\epsilon,h,j,m,\tilde{k})\right]\phi(x)\,\mathrm{d}x \\
&+\dot{V}_t(s,\eta_k,\epsilon,h,j,m,\tilde{k}) \\
&\quad\text{such that:} \\
&\quad V_t(s,\eta_k,\epsilon,h,j,m,\tilde{k}) \geq \max_{\tilde{j}\in\{m,r,a\}} V_t(s,\eta_k,\epsilon,\underline{h},\tilde{j},m,\tilde{k})
\end{aligned}$$

and symmetrically for households with state η_{-k}. $\phi(\cdot)$ denotes the pdf of a normal distribution with standard deviation σ_ϵ. I employ the methods outlined in Achdou et al. (2020) to solve this household problem.[43]

D.3 Short-run impact of falling relative price of ICT capital

What is the short-run effect of a falling relative price of ICT capital (a rise in q_{ict}) on wages and the interest rate? To develop an answer, note that the production function (1.2) can be written with capital expressed in units of the final (consumption) good using $\tilde{K}_{ict} = \frac{K_{ict}}{q_{ict}}$:

$$Y = F(N_m, N_r, N_a, q_{ict}\tilde{K}_{ict}, K_s)$$

In this section, let $\tilde{r}_{ict}$ denote the interest rate on price-adjusted ICT capital $\tilde{K}_{ict}$, such that $\tilde{r}_{ict} = q_{ict}F_{\tilde{K}_{ict}}(N_m, N_r, N_a, q_{ict}\tilde{K}_{ict}, K_s)$. The following proposition holds[44]

Proposition 4. *Consider the impact of a change in q_{ict} in the short term, assuming that factor inputs N_m, N_r, N_a, $\tilde{K}_{ict}$, and K_s do not change.*[45] *The impact of a change in q_{ict} on*

[43] To deal with the stopping-time nature of the problem I further rely on the note "Liquid and Illiquid Assets with Fixed Adjustment Costs" by Greg Kaplan, Peter Maxted and Benjamin Moll, accessed on November 6, 2021, at https://benjaminmoll.com/wp-content/uploads/2020/06/liquid_illiquid_numerical.pdf.

[44] The exposition and proofs follow along the lines of Acemoglu and Restrepo (2018a).

[45] The assumption here is therefore that K_{ict} adjusts so that $\tilde{K}_{ict}$ stays constant when q_{ict} changes.

wages and the rental rate of ICT capital is given by:

$$\frac{\mathrm{d}\ln w_i}{\mathrm{d}\ln q_{ict}} = \frac{1-s_i}{\epsilon_{\tilde{K}_{ict}N_i}} - \frac{s_j}{\epsilon_{\tilde{K}_{ict}N_j}} - \frac{s_h}{\epsilon_{\tilde{K}_{ict}N_h}} - \frac{s_{K_s}}{\epsilon_{\tilde{K}_{ict}K_s}} \tag{P.1a}$$

$$\frac{\mathrm{d}\ln \tilde{r}_{ict}}{\mathrm{d}\ln q_{ict}} = 1 - \frac{s_i}{\epsilon_{\tilde{K}_{ict}N_i}} - \frac{s_j}{\epsilon_{\tilde{K}_{ict}N_j}} - \frac{s_h}{\epsilon_{\tilde{K}_{ict}N_h}} - \frac{s_{K_s}}{\epsilon_{\tilde{K}_{ict}K_s}} \tag{P.1b}$$

where I define the elasticity of substitution of input factor X with ICT capital as

$$\epsilon_{\tilde{K}_{ict}X} = -\frac{\mathrm{d}\ln\left(q_{ict}\tilde{K}_{ict}/X\right)}{\mathrm{d}\ln\left(F_{\tilde{K}_{ict}}/F_X\right)} \tag{1.15}$$

and the shares of income as

$$s_i = \frac{w_i N_i}{Y}, \quad s_{\tilde{K}_{ict}} = \frac{\tilde{r}_{ict}\tilde{K}_{ict}}{Y}, \quad s_{K_s} = \frac{r_s K_s}{Y}\ .$$

Proof. Rewrite (1.15) as

$$\epsilon_{\tilde{K}_{ict}N_i}(\mathrm{d}\ln F_{\tilde{K}_{ict}} - \mathrm{d}\ln F_{N_i}) = -\left(\mathrm{d}\ln\left(q_{ict}\tilde{K}_{ict}\right) - \mathrm{d}\ln(N_i)\right)$$

Using the facts that in the short run $\mathrm{d}\ln\tilde{r} = \mathrm{d}\ln F_{\tilde{K}_{ict}} + \mathrm{d}\ln q_{ict}$, $\mathrm{d}\ln w_i = \mathrm{d}\ln F_{N_i}$, and $\mathrm{d}\ln\tilde{K}_{ict} = \mathrm{d}\ln K_s = \mathrm{d}\ln N_i = 0$, we have that

$$\begin{aligned} \epsilon_{\tilde{K}_{ict}N_i}(\mathrm{d}\ln\tilde{r}_{ict} - \mathrm{d}\ln q_{ict} - \mathrm{d}\ln F_{N_i}) &= -(\mathrm{d}\ln q_{ict}) \\ \Rightarrow \quad (\mathrm{d}\ln w_i - \mathrm{d}\ln\tilde{r}_{ict})\epsilon_{\tilde{K}_{ict}N_i} &= (1-\epsilon_{\tilde{K}_{ict}N_i})\,\mathrm{d}\ln q_{ict} \end{aligned} \tag{1.16}$$

We further have that because of constant returns to scale

$$\begin{aligned} Y &= w_a N_a + w_r N_r + w_m N_m + \tilde{r}_{ict}\tilde{K}_{ict} + r_s K_s \\ \mathrm{d}Y &= \mathrm{d}w_a N_a + \mathrm{d}w_r N_r + \mathrm{d}w_m N_m + \mathrm{d}\tilde{r}_{ict}\tilde{K}_{ict} + \mathrm{d}r_s K_s \quad |:Y \\ \frac{\mathrm{d}q_{ict}}{q_{ict}}\frac{q_{ict}F_{\tilde{K}_{ict}}\tilde{K}_{ict}}{Y} &= \mathrm{d}w_a\frac{s_{N_a}}{w_a} + \mathrm{d}w_r\frac{s_{N_r}}{w_r} + \mathrm{d}w_m\frac{s_{N_m}}{w_m} + \mathrm{d}\tilde{r}_{ict}\frac{s_{\tilde{K}_{ict}}}{\tilde{r}_{ict}} + \mathrm{d}r_s\frac{s_{K_s}}{r_s} \\ \mathrm{d}\ln q_{ict}\ s_{\tilde{K}_{ict}} &= \mathrm{d}\ln w_a\ s_{N_a} + \mathrm{d}\ln w_r\ s_{N_r} + \mathrm{d}\ln w_m\ s_{N_m} + \mathrm{d}\ln\tilde{r}_{ict}\ s_{\tilde{K}_{ict}} + \mathrm{d}\ln r_s s_{K_s} \end{aligned} \tag{1.17}$$

Combining (1.16) and (1.17) yields (P.1a) and (P.1b). □

Figure 1.19 reports the wage elasticities over time using the calibrated values of the production function and the factor inputs from the baseline transition. Since routine labor is highly complementary to ICT capital, its wage rate falls when q_{ict} rises. This stands in contrast to wages in manual and abstract occupations, which both perform work that is relatively easy to substitute with ICT capital. As can be seen, the magnitudes of the elasticities grow over time, as ICT capital becomes more important in production. For

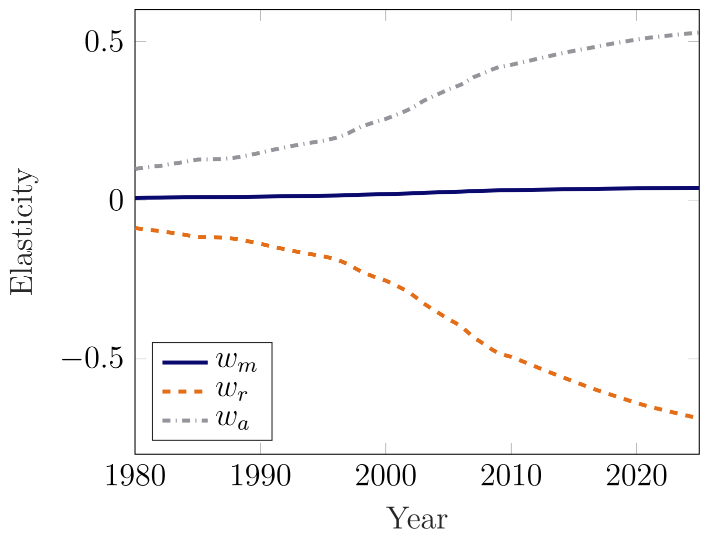


Figure 1.19: Short-run wage elasticities $\frac{\mathrm{d}\ln w_i}{\mathrm{d}\ln q_{ict}}$.

instance, a one percent increase in q_{ict} leads to a fall in w_r of 0.09 percent in 1980, and of 0.64 percent in 2020.

D.4 Computation of counterfactual densities

To compute the counterfactual densities in Section 4 I proceed as follows. I start with the distribution of households over the state space in 1980 $\Gamma_{1980}(i)$. Denote the mass of workers of skill type $\tilde{s}$ in occupation j in 1980 as follows:

$$g_{1980,in,j,\tilde{s}} = \int_{i:\{h=\underline{h}\wedge s=\tilde{s}\}} \mathbf{1}_{\{j^*_{1980}(i)=j\}} \mathrm{d}\Gamma_{1980}(i)$$
$$g_{1980,ex,j,\tilde{s}} = \int_{i:\{h=\bar{h}\wedge s=\tilde{s}\}} \mathbf{1}_{\{j^*_{1980}(i)=j\}} \mathrm{d}\Gamma_{1980}(i)$$

where $j^*_{1980}(i)$ denotes optimal occupational choices in 1980, and in and ex in the index of the densities abbreviates "inexperienced" and "experienced", respectively.

Note that when solving for the transition path I discretize time into periods (Kaplan, Moll, et al., 2018). For each point t on the discretized time grid I compute for each skill type $\tilde{s}$ the average probability across all inexperienced households to choose a certain occupation j, and denote this probability by $p_{t,\tilde{s},j}$, i.e.

$$p_{t,\tilde{s},j} = \frac{\int_{i:\{h=\underline{h}\wedge s=\tilde{s}\}} \mathbf{1}_{\{j^*_t=j\}} \mathrm{d}\Gamma_t(i)}{\int_{i:\{h=\underline{h}\wedge s=\tilde{s}\}} \mathrm{d}\Gamma_t(i)} \tag{1.18}$$

This effectively averages over the wealth (m and $\tilde{k}$) and idiosyncratic productivity (η and ϵ) dimensions, while conditioning on human capital and skill type. Γ_t refers to the distribution at time t during the baseline transition. Similarly, for $j \in \{m, r, a\}$, I compute the average probability $x_{t,\tilde{s},j}$ that experienced households exit their current occupation j and become inexperienced households

$$x_{t,\tilde{s},j} = \frac{\int_{i:\{j^*_{t-1}=j \wedge h=\bar{h} \wedge s=\tilde{s}\}} \mathbf{1}_{\{j^*_t \neq j\}} \mathrm{d}\Gamma_t(i)}{\int_{i:\{j^*_{t-1}=j \wedge h=\bar{h} \wedge s=\tilde{s}\}} \mathrm{d}\Gamma_t(i)} \tag{1.19}$$

I then use these probabilities to iterate forward the densities of skill type $\tilde{s}$ in the following way. For inexperienced households in occupation j:

$$\begin{aligned} g_{t+1,in,j,\tilde{s}} = p_{t,\tilde{s},j} \Bigg\{ & \underbrace{\sum_{k \in \{m,r,a\}} \mathrm{e}^{-\lambda^k_h \cdot \mathrm{dt}} g_{t,in,k,\tilde{s}}}_{\text{inexp. HHs who do not become experienced}} \\ & + \underbrace{\sum_{k \in \{m,r,a\}} \left[x_{t,\tilde{s},k} \left(\mathrm{e}^{-\zeta \cdot \mathrm{dt}} g_{t,ex,k,\tilde{s}} + (1 - \mathrm{e}^{-\lambda^k_h \cdot \mathrm{dt}}) g_{t,in,k,\tilde{s}} \right) \right]}_{\text{exp. HHs who exit their occ.}} \\ & + \underbrace{(1 - \mathrm{e}^{-\zeta \cdot \mathrm{dt}}) \sum_{k \in \{m,r,a\}} g_{t,ex,k,\tilde{s}}}_{\text{dead exp. HHs}} \Bigg\} \end{aligned} \tag{1.20}$$

And for experienced households in occupation j:

$$g_{t+1,ex,j,\tilde{s}} = (1 - x_{t,\tilde{s},j}) \Big\{ \mathrm{e}^{-\zeta \cdot \mathrm{dt}} g_{t,ex,j,\tilde{s}} + (1 - \mathrm{e}^{-\lambda^j_h \cdot \mathrm{dt}}) g_{t,in,j,\tilde{s}} \Big\} \tag{1.21}$$

The last step is to aggregate these densities across all skill types s.

Figure 1.20 plots both the actual mass of households in the routine occupations (i.e. using Γ_t), as well as the one obtained from the iterative procedure described here. It reveals that by using the iterative procedure I recover the actual employment share during the transition very well.

Figure 1.5 For this counterfactual, in which I assume that only newborns can make occupational choices, I first set all exit probabilities $x_{t,s,j} = 0$ in (1.20) and (1.21). Furthermore, only newborns can choose a new occupation, i.e. (1.20) becomes

$$\begin{aligned} g_{t+1,in,j,\tilde{s}} = {} & \mathrm{e}^{-\zeta \cdot \mathrm{dt}} \mathrm{e}^{-\lambda^j_h \cdot \mathrm{dt}} g_{t,in,j,\tilde{s}} \\ & + p_{t,\tilde{s},j} (1 - \mathrm{e}^{-\zeta \cdot \mathrm{dt}}) \sum_{k \in \{m,r,a\}} \left[g_{t,ex,k,\tilde{s}} + \mathrm{e}^{-\lambda^k_h \cdot \mathrm{dt}} g_{t,in,k,\tilde{s}} \right] . \end{aligned}$$

Figure 1.7 To arrive at the counterfactual employment shares that use policy functions of the liquid-wealthy households only, I replace $\Gamma_t(i)$ in (1.18) and (1.19) with

$$\Gamma_t(i) \cdot \mathbf{1}_{i:\{m \geq 70.(90.)\ \text{prctl.}\}} .$$

Given that households with high liquid assets have systematically different (i.e. higher) productivity draws than the average population, I correct for this as follows. For each t, I

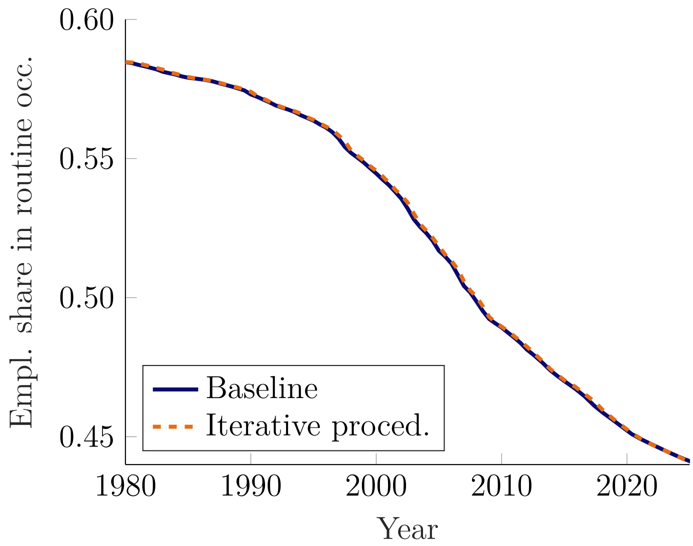


Figure 1.20: Employment share in the routine occupation.

Notes: Blue solid: true employment share during the transition ($\int_{i:j_t^*=r} \mathrm{d}\Gamma_t(i)$). Orange dashed: employment share constructed from iterative procedure ($\sum_s g_{t,in,r,s} + g_{t,ex,r,s}$).

compute (1.18) and (1.19) for each point on the discretized grids of the two productivity processes. I then weight these probabilities according to their occurrence in the whole population (i.e. I use the invariant stationary distribution for weighting) to arrive at aggregate values for $p_{t,\tilde{s},j}$ and $x_{t,\tilde{s},j}$. I then iterate on (1.20) and (1.21) as described above.

E Policies Addressing Switchers to Manual Occupation

In Section 6 the policies were targeted only at experienced routine workers who switched to the abstract occupation. In this section, I assume instead that experienced workers switching to the manual occupation become eligible to the loan (wage replacement).

Figure 1.21 plots aggregate welfare when introducing the policies minus welfare in the baseline transition (ΔW). As becomes clear, implementing the programs for switchers to manual does not increase social welfare. Figure 1.22 reveals why this is the case, by decomposing welfare changes over time for $L = 10,000$\$ and $R = 420$\$, i.e. the programs that were optimal when targeting them at switchers to abstract. While at least in the case of the wage replacement program the partial equilibrium effect on welfare is positive (orange dashed line), the general equilibrium effects of both programs on welfare are negative. This is because the additional routine workers that switch to the manual occupation put downward pressure on the manual wage w_m along the transition path relative to the baseline. This harms low-skilled and therefore on average poor households and therefore decreases welfare.

Apart from these general equilibrium effects, however, there are further reasons why one should not expect programs alleviating the costs of experienced routine workers switching

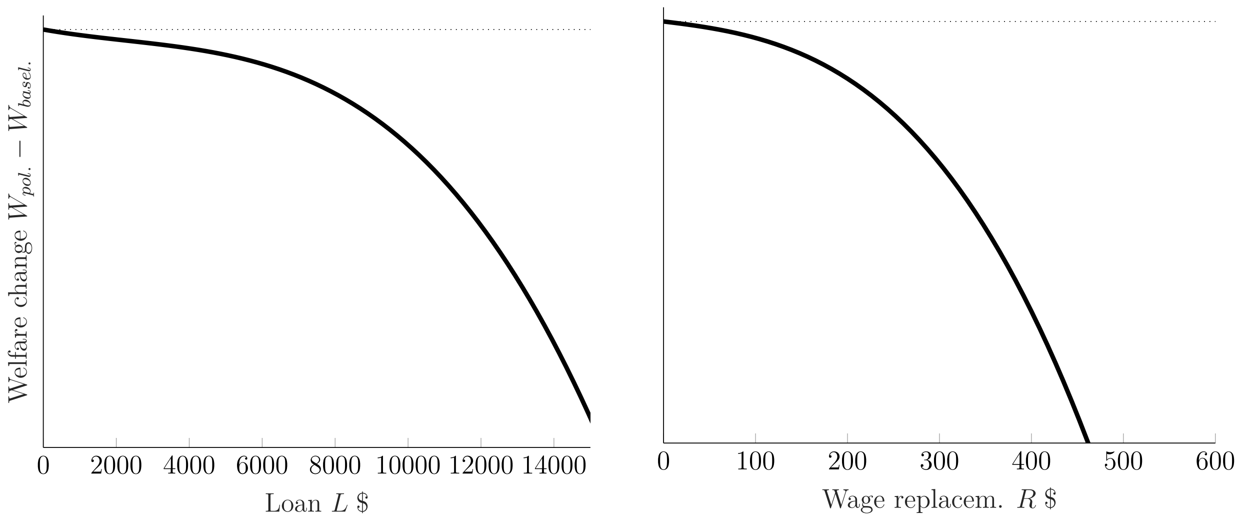


Figure 1.21: Absolute changes in welfare under the policies (targeted at switchers from routine to manual) compared to the baseline.

Notes: Nominal quantities are expressed in 2017 US dollars.

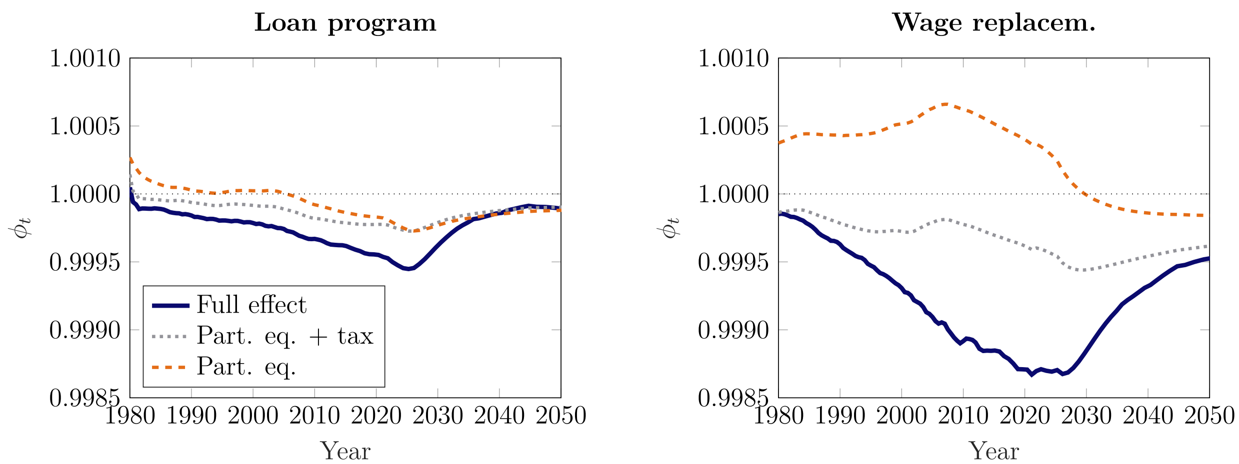


Figure 1.22: Changes in welfare under the policies (targeted at switchers from routine to manual) compared to the baseline over time.

Notes: Left: loan program with $L = 10,000\$$. Right: wage replacement program with $R = 420\$$. ϕ_t is s.t. $\int_i u(\phi_t c_t^{basel}, \ell_t^{basel})\, \mathrm{d}\Gamma_t^{basel}(i) = \int_i u(c_t^{pol}, \ell_t^{pol})\, \mathrm{d}\Gamma_t^{pol}(i)$ holds. "Part. eq." holds fixed taxes $\tau_{l,t}$ and factor prices at their values from the baseline transition without policy. "Part. eq. + tax" holds fixed factor prices only. "Full effect" is the full general equilibrium effect.

to the manual occupation to provide many benefits. First, the manual wage rises only marginally relative to the routine wage, especially when one contrasts this to the rise in the abstract to routine wage (see Figure 1.23 below). It is therefore rarely the case that foregoing their occupation-specific human capital in the routine occupation by switching to the manual occupation is optimal for experienced routine workers. Second, the calibrated occupation-specific slopes a_j in Figure 1.2 imply that many medium-skilled workers bear much lower wage losses when switching to the manual than when switching to the abstract occupation ($a_r - a_m < a_a - a_r$). The elasticity of the manual labor supply with respect to the routine wage is therefore much higher than that of the abstract labor supply. For both of these reasons, the increase in the manual employment share, which is small quantitatively in any case, mostly runs via newborn or inexperienced households moving into manual occupations and less via experienced routine workers switching there. The policies also do not increase output on average, as was the case when targeting them at switchers to the abstract occupation (graph not shown).

F Additional Graphs and Tables

Table 1.12: Parameters and targets for the calibration of Φ_ϵ.

Moment	Data	Model
Variance: ann. log earnings	0.70	0.70
Variance: 1-year change	0.26	0.26
Kurtosis: 1-year change	14.9	15.0
Frac. 1-year log change < 0.2	66.5%	70.2%
Frac. 1-year log change > 1.0	6.6%	6.6%

Notes: Empirical moments (Data) are taken from Guvenen et al. (2021), except for the variance of annual log earnings, which is not reported there. Instead, I use the value reported in Kaplan, Moll, et al. (2018).

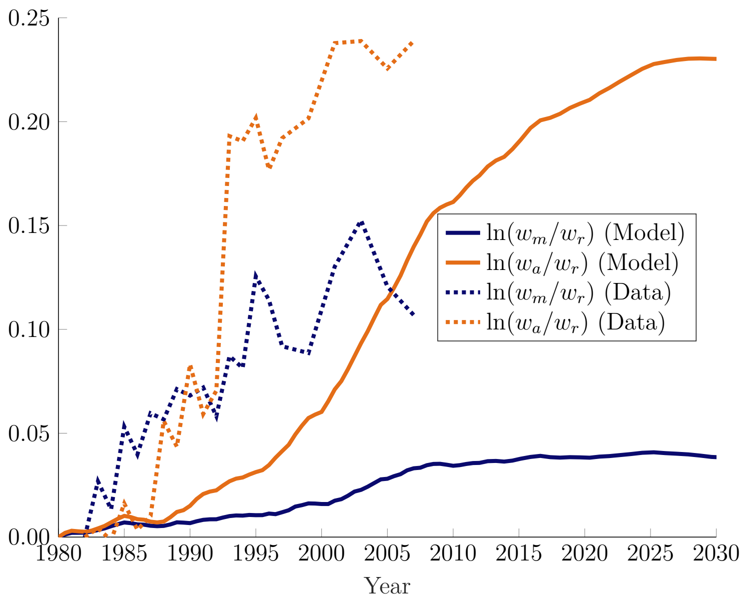


Figure 1.23: Model-implied change in relative wages per efficiency unit compared to 1980.
Notes: Empirical estimates are taken from the lower left panel of Figure 6 in Cortes (2016), in which he allows for heterogeneous tenure profiles, as I do in the model.

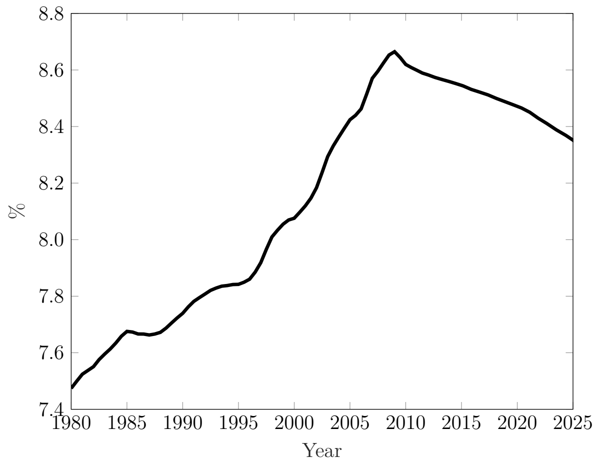


Figure 1.24: Model-implied (annual) interest rate r (in %).

Table 1.13: Hand-to-mouth shares by occupation.

	SCF (1989)			Model (init. steady state)		
	HtM	wealthy HtM	poor HtM	HtM	wealthy HtM	poor HtM
Manual	42%	20%	22%	41%	26%	15%
Routine	32%	20%	13%	34%	23%	11%
Abstract	16%	13%	3%	17%	12%	6%

Notes: Wealthy HtM are hand-to-mouth households with positive illiquid assets, poor HtM are those with zero illiquid assets.

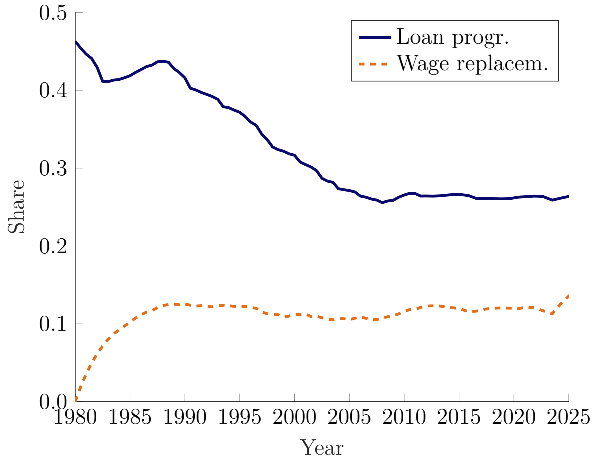


Figure 1.25: Share of eligible workers who work in the routine occupation under optimal programs.

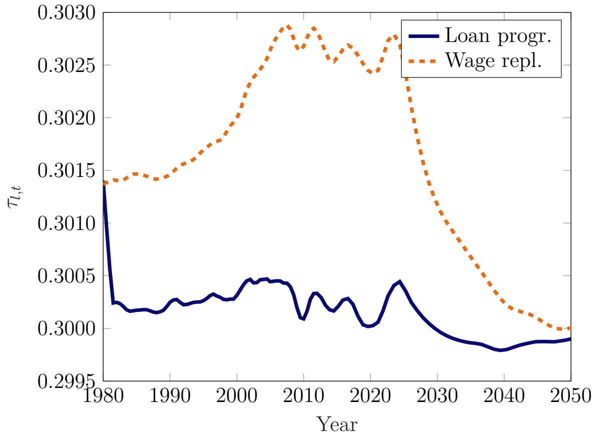


Figure 1.26: Labor income tax $\tau_{l,t}$ necessary to finance the policies.

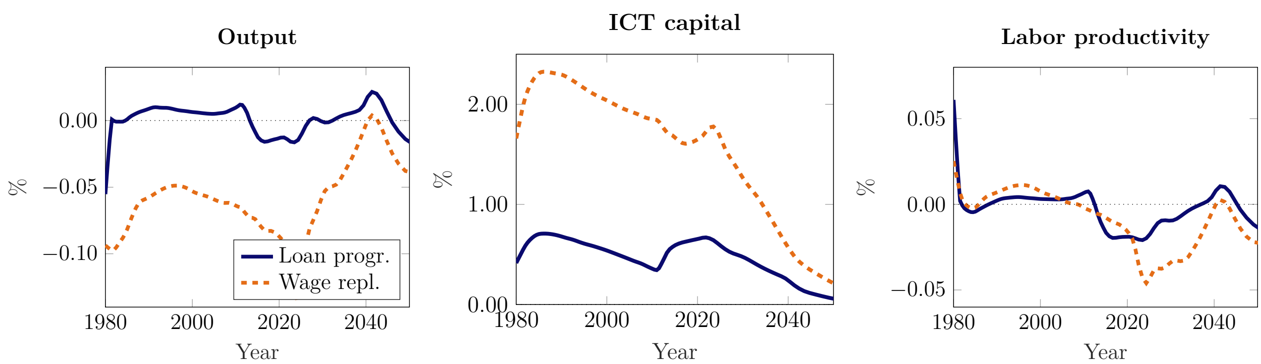


Figure 1.27: Deviations of variables when policies are introduced in steady state relative to steady state without policies.

Notes: Left: Output Y_t. Middle: ICT capital $K_{ict,t}$. Right: Labor prod. $Y_t / \left(\int_i \ell_t^* \, \mathrm{d}\Gamma_t(i)\right)$. Size of L and R are 10,000$ and 420$ respectively.

Chapter 2

Does Wealth Inequality Affect the Transmission of Monetary Policy?

Joint with Alexander Matusche.

Disclaimer: This chapter uses data from the Eurosystem Household Finance and Consumption Survey. The results published and the related observations and analysis may not correspond to results or analysis of the data producers.

1 Introduction

In the US, but also in other advanced economies, wealth inequality has risen considerably since the mid-1980s. In general, wealth inequality displays significant variation both over time as well as across space.[1] At the same time, recent advances in macroeconomics have led to a new discussion about the role of inequality for macroeconomic outcomes. In particular, several models predict that the wealth distribution is important for understanding the transmission of monetary policy.[2] We contribute to this debate by documenting empirically that wealth inequality affects the strength of the monetary transmission mechanism to the real economy. We provide three separate pieces of evidence, each showing that monetary policy has larger real effects when wealth inequality in the population is high. To the best of our knowledge, we are the first to offer a detailed account of this relationship.

Macroeconomists have long neglected wealth inequality when studying monetary policy. This was largely due to the complexity and computational demands of structural models that feature a non-degenerate wealth distribution. However, quantitative research also supported the view that wealth inequality had little importance for macroeconomic dynamics (Krusell

[1]See, for instance, Hubmer et al. (2021) for the evolution of wealth inequality in the US over the past century, Piketty (2014) for a detailed account of worldwide inequality, and the graphs and statistics we provide in this study.

[2]Examples are Auclert (2019), Bilbiie (2020), Gornemann et al. (2016), Kaplan, Moll, et al. (2018), Luetticke (2021), and Werning (2015)

and Smith, 1998). Recently, however, this view has been challenged, and the question of how "inequality matters for macro" (Ahn et al., 2018) has received new attention.

Several mechanisms have been proposed through which the distribution of wealth affects the transmission of monetary policy. Building on the observation that wealthier households are more likely to invest in stocks, Melcangi and Sterk (2020) argue that, when the rich hold a greater share of wealth, an interest rate cut leads to a stronger rebalancing towards equities which results in a stock market and investment boom. Luetticke (2021) finds that greater wealth inequality mutes the response of aggregate investment and amplifies that of consumption, with the overall effect on output approximately canceling out. The reason is that wealthy households with high marginal propensities to invest benefit from contractionary monetary policy, whereas incomes of asset-poor households with high marginal propensities to consume fall. Kaplan, Moll, et al. (2018) emphasize the role of hand-to-mouth behavior. Households who are close to the borrowing constraint and hold few (liquid) assets have high marginal propensities to consume such that monetary policy has stronger effects when there are more hand-to-mouth households. In this paper we ask whether there is a state dependence of monetary policy transmission on wealth inequality in the data that would validate the predictions of these studies.

We tackle the question of how wealth inequality alters the transmission of monetary policy in three steps. We first use aggregate time series data, separately for the US and the UK, and ask whether the effects of monetary policy are state-dependent. To this end, we estimate state-dependent local projections, as suggested by Auerbach and Gorodnichenko (2013), using the top 10% wealth share as a state variable. In the US and the UK, we find that in regimes of high inequality monetary policy has historically had a larger impact on real activity than in regimes of low inequality. The differences are quantitatively important. In the US, while industrial production contracted by up to 2 percent in times of high wealth inequality in response to a 25 basis points (bp) increase in the federal funds rate, it shows no statistically significant contraction in times of low inequality. In the UK, responses of output and unemployment to monetary policy are smaller overall but, as in the US, they are relatively larger in times of high wealth inequality.

While the appeal of our first approach is its simplicity and the availability of high-quality data at the aggregate level, the drawback is that we cannot rule out that other variables that have co-moved with wealth inequality over time are responsible for our results. Therefore, we turn to cross-sectional analyses in the second part of our study, which allows us to control for confounding factors. We use estimates provided by the Internal Revenue Service (IRS) on total wealth held by the richest households in each US state to construct measures of state-level wealth inequality. In line with the results on the aggregate level, we find that both output and unemployment in US states that display more wealth inequality react more strongly to common monetary policy shocks. This relationship has been especially strong since the early 1980s. In a third step, we construct measures of wealth inequality for Euro Area countries using data from the ECB's Household Finance and Consumption Survey

(HFCS). We find that Euro Area countries with high levels of inequality react more strongly to common monetary policy shocks.

Based on the consistent findings in all three settings and after conducting various robustness checks, we conclude that higher wealth inequality is correlated with a stronger transmission of interest rate changes to the real economy. The strength of this correlation differs in the set-ups we study. Regarding the size of the output response, we estimate the largest effect on the US aggregate level. Here, an increase in the top 10% share by one percentage point is associated with a 0.26 percentage points stronger average contraction in industrial production over the first three years after a 25 bp monetary policy shock. The effect is an order of magnitude smaller in the UK and in the cross-section of US states, and it is about 0.05 in the Euro Area countries.[3] Likewise, the effect of an increase in inequality on the average response of the unemployment rate ranges from 0.005 percentage points in our preferred specification for the US states to 0.018 on the US aggregate level and 0.041 in the Euro Area. These numbers point to an economically meaningful effect of wealth inequality on the transmission of monetary policy to the real economy.

The reduced-form evidence that we provide in this study can inform and discipline the design of future structural models that analyze monetary policy in the framework of heterogeneous agent models. It is also relevant for the current debate among policymakers about when and by how much to raise interest rates. Given the high levels of wealth inequality observed in many developed countries at the moment, our results indicate that a rate increase would have relatively strong effects on the real economy.

The remainder of our paper is structured as follows. Below, we review other related literature. We study the state-dependent effects of monetary policy on the aggregate level in Section 2. In Section 3 we analyze the cross-section of US states, and in Section 4 we turn to a cross-section of Euro Area countries. In Section 5 we conclude.

Related Literature We follow a long tradition of studies that have analyzed the effects of monetary policy empirically using time series methods on an aggregate level (Christiano et al., 1999, 2005; Coibion, 2012; Ramey, 2016). In particular, in Section 2 we use state-dependent local projections to estimate the dependence of monetary policy transmission on the distribution of wealth. This approach has been advocated by Auerbach and Gorodnichenko (2013) and Ramey and Zubairy (2018), who ask whether government multipliers are larger or smaller during times of economic slack. In the context of monetary policy, state-dependent local projections have been used by Alpanda and Zubairy (2019), Ascari and Haber (2021), and Tenreyro and Thwaites (2016). Most related to our study is Alpanda and Zubairy (2019), who find that high levels of household debt mute the effects of interest rate changes. We corroborate this finding in Section 4, but show that wealth inequality influences the strength of monetary policy transmission even when controlling for the debt-

[3] At the US state level, due to data limitations, we only estimate the top 1% wealth share and its impact on monetary policy transmission, not the 10% wealth share as in the other sections of this study.

to-GDP ratio. Lastly, the study of Brinca et al. (2016) is close to ours. They ask how the size of fiscal multipliers depends on wealth inequality and find that higher inequality is associated with larger multipliers.

Our analysis of US states in Section 3 is similar to Carlino and DeFina (1998, 1999) and Owyang and Wall (2009) whose findings suggest a role for the industrial composition within a state as well as the presence of a large share of small firms in affecting the effectiveness of monetary policy. We add wealth inequality as an additional explanatory variable, which we construct based on estate tax returns.[4] Our analysis on Euro Area countries in Section 4 is related to Almgren et al. (2020) who find that countries in which many households hold few liquid assets react more strongly to monetary policy. While we can validate their results in our set-up, we take a somewhat broader perspective by focusing on wealth inequality, and find a significant effect of wealth inequality even when controlling for other explanatory variables. Slacalek et al. (2020) also investigate the effects of monetary policy shocks in the Euro Area but have a different focus than we do. They only study the four largest economies and decompose their responses to monetary policy shocks using structural models. In contrast, we look at all Euro Area countries and keep to a reduced-form analysis without explicitly quantifying different theoretical transmission channels.

We are not aware of empirical studies that analyze the dependence of monetary policy transmission on wealth inequality, but several authors have investigated the reverse question, i.e., to what extent monetary policy affects inequality. For instance, Coibion et al. (2017) find that expansionary monetary policy tends to lower inequality in income and consumption, though they do not study wealth inequality as an outcome variable because of data limitations. Adam and Zhu (2016) draw similar conclusions using Euro Area survey data from the HFCS. In contrast, Andersen et al. (2021) who use administrative household-level data from Denmark, find that expansionary monetary policy increases inequality.

2 Evidence from State-Dependent Local Projections

2.1 United States

We obtain measures of wealth inequality for the US from Piketty et al. (2018), made publicly available via the World Inequality Database (WID). Piketty et al. (2018) follow the approach in Saez and Zucman (2016) to estimate the wealth distribution. They define household wealth as the current market value of all the assets owned by the household net of all its debt and estimate household wealth primarily by capitalizing income tax data.[5]

As our benchmark measure of inequality we use the wealth share held by the top 10%

[4]Estate tax returns have been used before to construct measures of wealth inequality at the US aggregate level (Kopczuk, 2015; Kopczuk and Saez, 2004; Saez and Zucman, 2016), however not at the state level to the best of our knowledge.

[5]See Saez and Zucman (2016) for a detailed discussion of the definition of wealth used and the capitalization approach.

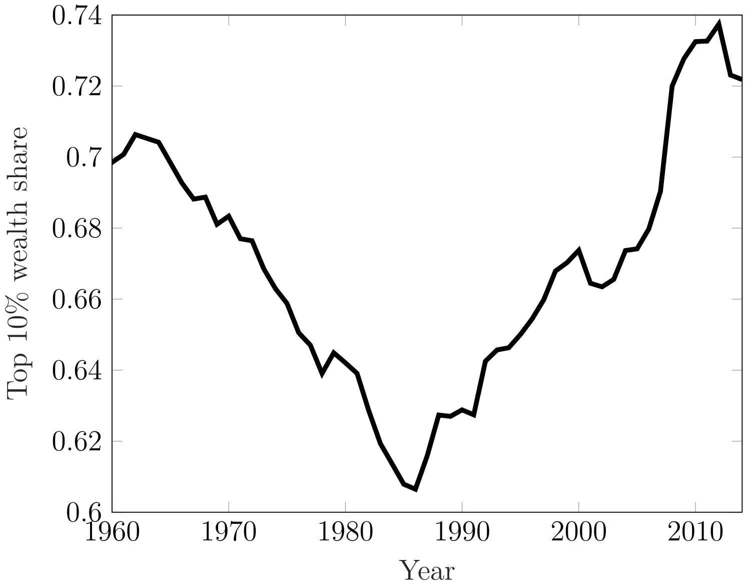


Figure 2.1: Top 10% wealth share in the US.

Notes: Data obtained from the WID (see text).

of the wealth distribution. The 90th percentile of the wealth distribution roughly separates the "wealth middle class" and the very wealthy and is therefore a statistic commonly used to assess the degree of wealth inequality in a given population (Alvaredo, Chancel, et al., 2018). We show that our results are robust to taking other measures, such as the top 1% share and the Gini index, at the end of this section. Figure 2.1 shows the u-shaped evolution of the wealth share of the richest 10% over time. It peaked in 1962 at 71%, declined to 61% until 1986, and rose from then on, reaching 74% in 2012.

In this section we use time series data on US aggregate variables in order to assess how wealth inequality among households affects the transmission of monetary policy.[6] We estimate local projections, which have been proposed by Jordà (2005) as a way of directly estimating impulse response functions (IRFs) to identified shocks. Recent empirical studies have relied heavily on this approach (Auerbach and Gorodnichenko, 2013; Coibion et al., 2017; Ramey, 2016). To be precise, for each horizon $h \in [0, 1, \ldots, H]$, we estimate the following model:

$$\begin{aligned} y_{t+h} = \alpha_h + \beta_h \cdot i_t + \beta_h^+ \cdot (\text{ineq}_t \cdot i_t) \\ + \sum_{p=1}^{P} \Gamma_{h,p} \cdot X_{t-p} + \sum_{p=1}^{P} \Gamma_{h,p}^+ \cdot (\text{ineq}_t \cdot X_{t-p}) + u_{t+h} \end{aligned} \tag{2.1}$$

where y_t is an outcome variable of interest (e.g. the log of industrial production), i_t is the federal funds rate, ineq_t is our measure of wealth inequality, X_t is a vector of control variables and u_t is an error term with $\mathbb{E}[u_t] = 0$.

The non-standard elements in equation (2.1) are the interaction terms between inequal-

[6]The aggregate time series data for the US and the UK are taken from the Federal Reserve Economic Database (FRED).

ity and X_t. We introduce them to capture the potential state dependence between wealth inequality and the effects that monetary policy has on the economy.[7] In times of low inequality, the response of y_t to a change in the federal funds rate is primarily governed by the parameters β_h. As inequality rises, the parameters β_h^+ become increasingly important in determining the response of y_t. Hence, conditional on a specific value of inequality at time t, $\left(\beta_h + \beta_h^+ \cdot \text{ineq}_t\right)$ gives the response of y to a 100 basis points increase in the interest rate h months after it occurred.

To make our results comparable to the empirical literature on monetary policy shocks, we follow Ramey (2016) in designing our baseline specification. In particular, we use data at monthly frequency, α_h collects a constant and a linear time trend, we set $P = 2$, and we include as controls in X_t the log of industrial production, the unemployment rate, the log of the consumer price index and the commodity price index, as well as the federal funds rate.[8] As our inequality measure is only available at annual frequency, we assign each month of a given year that year's observation. Since the top 10% share moves relatively slowly, other forms of interpolation lead to virtually the same results. We estimate IRFs up to a horizon of $H = 36$ months.

We instrument the federal funds rate with an exogenously identified monetary policy shock series, i.e. we estimate IV local projections as proposed by Stock and Watson (2018). The shock series we employ are the narratively identified monetary policy shocks from C. Romer and D. Romer (2004), which have been further extended by Coibion et al. (2017). C. Romer and D. Romer (2004) use narrative and quantitative records of meetings of the Federal Open Market Committee (FOMC) to derive a measure of intended changes in the federal funds rate around FOMC meetings. They then regress this measure on the Federal Reserve's internal forecasts of several aggregate variables. The residuals of this projection are then used as monetary policy shocks as they are relatively free of endogenous responses of the central bank to the state of the economy. As an instrument for the interaction term between inequality and the interest rate we use the product of the contemporaneous Romer & Romer shock and previous year's level of inequality, i.e. we lag inequality by one year. This ensures that no contemporaneous correlation between the shocks and inequality biases the estimates of β_h^+.[9] The major advantage of this shock series for our purposes is that it begins as early as 1969m3 such that our sample encompasses two time periods of relatively

[7]Instead of assuming a simple linear interaction term we have also experimented with more complicated, so-called "transition functions" $F(\text{ineq}_t)$, that map the degree of inequality into a value between zero and one (see for instance Auerbach and Gorodnichenko, 2013). The results were very similar to the ones shown here, which is why we opt for the more parsimonious specification using the linear interaction term, i.e. we set $F(\text{ineq}_t) = \text{ineq}_t$.

[8]Excluding the linear time trend only marginally affects the results. We do not include contemporaneous variables of the controls in X_t, i.e., using the terminology of Ramey (2016), we do not make a recursiveness assumption. At the end of this section we show that our results are robust to assuming recursivity.

[9]The results of Coibion et al. (2017) suggest such a correlation between monetary policy and income inequality. They do not consider wealth inequality as an outcome variable, however. Instrumenting the interaction term with both contemporaneous Romer & Romer shocks and levels of inequality instead does not affect our results. See also Appendix A.1.

high wealth inequality (the 1960s and the 2000s), and a period of relatively low inequality in the 1980s (see Figure 2.1).

Auerbach and Gorodnichenko (2013) stress that local projections offer important advantages in the context of modeling state dependence relative to other time series methods, such as Smooth Transition Vector Autoregressions (STVAR). Most importantly, when using local projections to produce impulse responses to a monetary policy shock for a given value of inequality at time t, one does not have to assume a path for inequality over the considered horizon. Rather, an IRF at horizon h gives the *average* response of y_{t+h} to a monetary policy shock that occurred at an inequality level of ineq_t. For instance, a systematic, endogenous response of wealth inequality to monetary policy as suggested by the results in Coibion et al. (2017), is not ruled out when estimating impulse responses. This stands in stark contrast to the IRFs constructed from an STVAR, where one has to make an assumption on how inequality evolves over the considered horizon of the IRF.

Figure 2.2 shows the estimated IRF for industrial production.[10] Here and throughout this paper we consider the effects of an interest rate increase of 25 basis points. Focus first on the dashed black line in Figure 2.2. It depicts the unconditional response of industrial production to a monetary policy shock. We obtain it by estimating equation (2.1), but leaving out the interaction terms with inequality, i.e. we impose that there is no state dependence on wealth inequality. We obtain the familiar result that a contractionary monetary policy shock depresses real activity in the economy, especially so at horizons between one and two years. The shape and the magnitude of the IRF are well in line with previous empirical results (Ramey, 2016).

Next we allow the effects of monetary policy to be state-dependent, i.e. we reintroduce the inequality interaction terms in equation (2.1). The blue line depicts the IRF when a monetary policy shock hits the economy during a regime of low inequality, the red line when it hits during a regime of high inequality. Low and high inequality refer to the first and third quartile of observed wealth inequality values in our sample (a top 10% wealth share of 62.9% and 67.0% respectively). The shaded areas indicate 90% confidence intervals constructed from Newey and West (1987) standard errors.

The core finding is that the IRFs in the two regimes are markedly different. On the one hand, the real effects of interest rate movements were strong during times of high inequality, leading to a contraction of industrial production by up to 2.1%. During times of low inequality, on the other hand, real activity was hardly affected by monetary policy. At short horizons we even obtain significant positive responses to a contractionary shock.[11] The IRF then approaches zero and later becomes negative, however never becomes statistically significant. The average response of industrial production over all estimated horizons differs between the high and the low inequality regime by 1.1 percentage points. Given the

[10] IRFs for other variables can be found in Figure 2.12 in Appendix A.

[11] The positive response at short horizons owes to some extent to the fact that we do not make any recursivity assumption in equation (2.1). As we show in the robustness checks, imposing recursivity leads to negative or insignificant responses at short horizons even in a regime of low inequality.

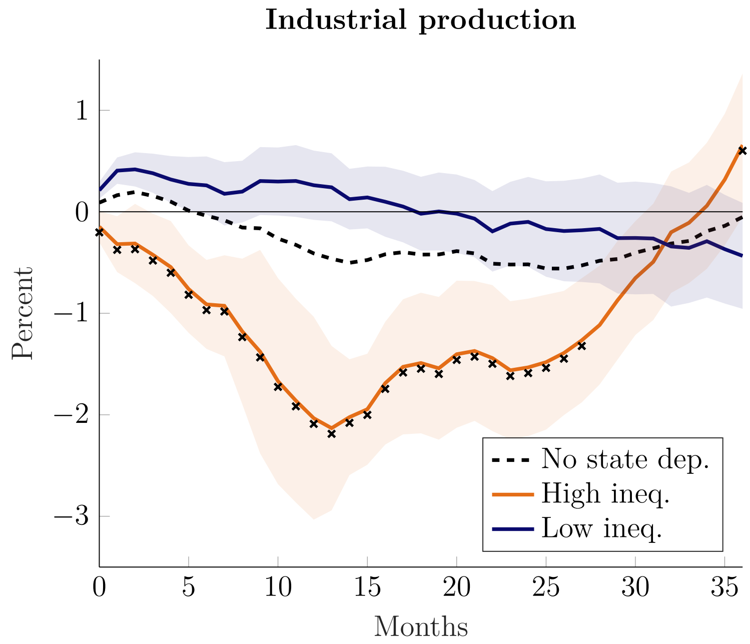


Figure 2.2: IRF of industrial production to 25 bp change of the fed. funds rate.
Notes: Black dashed: No state dependence is imposed. Blue: Regime of low inequality (first quartile of observed inequality). Red: Regime of high inequality (third quartile). Confidence intervals are at the 90% level and based on Newey and West (1987) standard errors. Black crosses indicate horizons at which we reject $H_0 : \beta_h^+ = 0$ at a 90% confidence level.

difference in the top 10% wealth share between the two regimes of 4.1 percentage points, an increase in this measure of inequality by one percentage point is associated with a 0.26 percentage points larger average contraction of industrial production in the first three years.

We can formally test for the state dependence of monetary policy by conducting a series of t-tests on the parameters β_h^+ for $h \in [0, 1, \ldots, H]$. The results of these tests are indicated by the black crosses in Figure 2.2. At the indicated horizons, we reject the null hypothesis that $\beta_h^+ = 0$ at a 90% confidence level. As can be seen, at all horizons during the first two years after the shock we can reject the hypothesis of no state dependence.

Figure 2.3 shows the reaction of the unemployment rate to a monetary policy shock. A similar pattern emerges as for industrial production. While unemployment barely rises after the shock under low inequality, there is a pronounced spike in unemployment if the shock takes place in a regime of high wealth inequality. Moreover, we find the state dependence of the IRFs to be significant at several horizons in the first and second year after the shock. Using the same back-of-the-envelope calculation as before, our results indicate that the unemployment response becomes on average 0.02 percentage points larger when the top 10% wealth share grows by one percentage point.

As was the case also for industrial production, at very high horizons the relationship between the two regimes switches. One potential explanation for this pattern is that at late horizons the local projections pick up the beginning of the next phase of the business cycle. This would account for a more pronounced boom at large horizons under high inequality after the shock caused a more pronounced recession at shorter horizons.

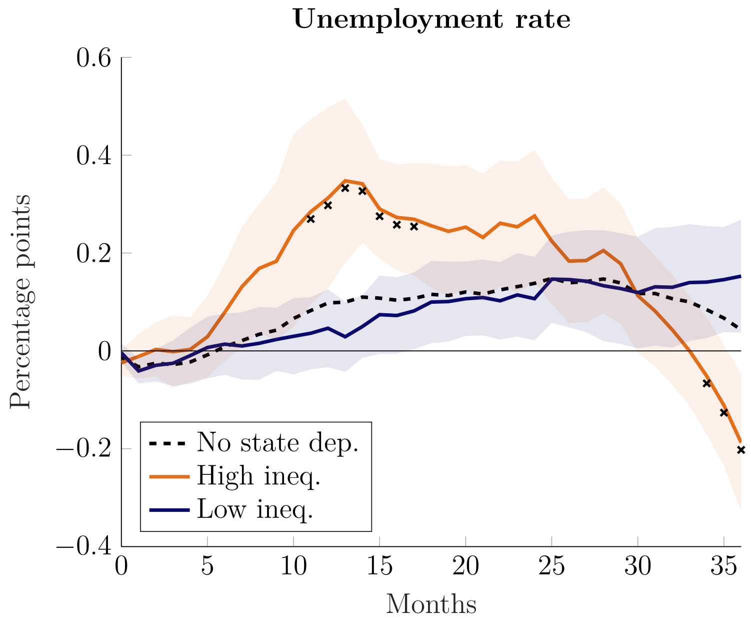


Figure 2.3: IRF of the unemployment rate to 25 bp change of the fed. funds rate.

Notes: Black dashed: No state dependence is imposed. Blue: Regime of low inequality (first quartile of observed inequality). Red: Regime of high inequality (third quartile). Confidence intervals are at the 90% level and based on Newey and West (1987) standard errors. Black crosses indicate horizons at which we reject $H_0 : \beta_h^+ = 0$ at a 90% confidence level.

Could it be that the endogenous reaction of the monetary authority is driving the differential results? For instance, the Federal Reserve might lower its rate faster in times of low inequality in response to a contractionary shock, thereby stimulating real activity. Figure 2.4 shows that this is not the case. It depicts the IRF of the federal funds rate. While over the first half year the responses are not statistically different from each other in the two regimes, at later horizons the reaction of the Federal Reserve becomes more accommodating in a regime of high inequality.

Robustness We assess the robustness of our results in several ways. For brevity, we show the corresponding IRFs of industrial production in Appendix A. First, we use different measures of inequality, namely the Gini coefficient and the top 1% wealth share. While the results using the Gini index are nearly identical to the baseline results (Figure 2.13), when we use the top 1% share (Figure 2.14) the absolute differences between the two regimes' IRFs become even larger.

Second, traditional identification schemes for monetary policy shocks based on timing restrictions have assumed that industrial production, unemployment and the price level do not respond contemporaneously to changes in the federal funds rate (see, for instance, Christiano et al., 1999). We can impose this recursivity by including contemporaneous values of log industrial production, unemployment and the log of the price indices in equation (2.1) (Ramey, 2016). As Figure 2.15 shows, assuming recursivity dampens the positive response of industrial production at low horizons but it remains the case that the response is stronger when inequality is high.

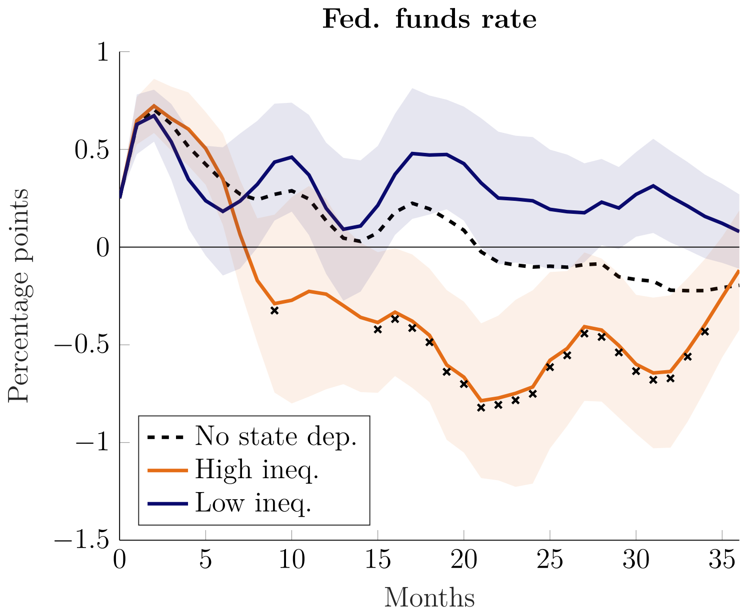


Figure 2.4: IRF of fed. funds rate to 25 bp change of the fed. funds rate.
Notes: Black dashed: No state dependence is imposed. Blue: Regime of low inequality (first quartile of observed inequality). Red: Regime of high inequality (third quartile). Confidence intervals are at the 90% level and based on Newey and West (1987) standard errors. Black crosses indicate horizons at which we reject $H_0 : \beta_h^+ = 0$ at a 90% confidence level.

Third, Coibion (2012) argues that IRFs to monetary policy shocks derived from Romer & Romer shocks are sensitive to the inclusion of the early 1980s. During this time period the Federal Reserve targeted non-borrowed reserves and the federal funds rate was a less suitable measure of the monetary policy stance. We check the robustness of our results by excluding this time period, i.e. we use only the years 1984 to 2007 in our sample. As Figure 2.16 shows, our results qualitatively still hold for this subperiod, even though the dependence on wealth inequality is statistically significant at fewer horizons.

Lastly, we also consider the Chicago Fed National Activity Index (CFNAI) instead of log industrial production as the measure of real economic activity. The results are shown in Figure 2.17. While the results are not as stark as for industrial production, there are still several horizons in the first year at which the response is significantly stronger during times of high inequality. At very long horizons, however, the relationship reverses, which to some extent was also the case when using industrial production as the measure of economic activity.

2.2 United Kingdom

The extent to which we can apply our method to other countries than the US is limited by the availability of long enough time series data, in particular on wealth inequality. However, we are able to corroborate our findings using data from the United Kingdom. The WID provides data taken from Alvaredo, Atkinson, et al. (2016) on the top 10% wealth share. We plot it in Figure 2.5. There are some similarities to the dynamics in the US but while

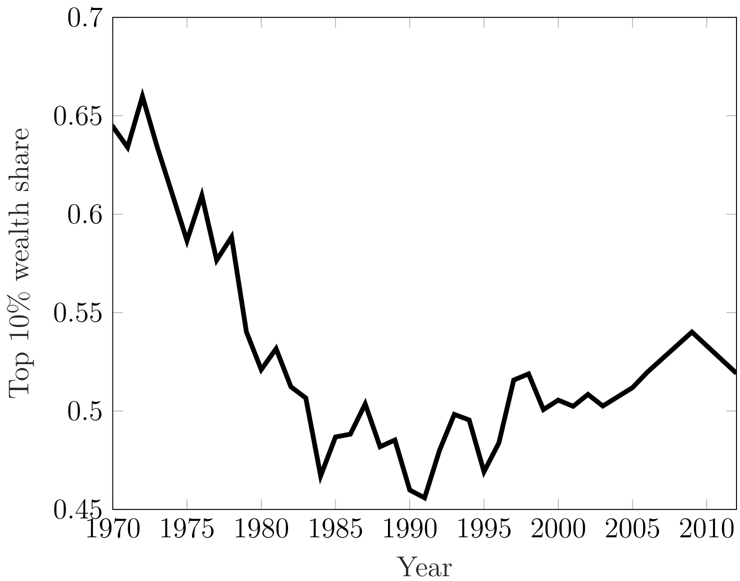


Figure 2.5: Top 10% wealth share in the UK.
Notes: Data obtained from the WID (see text).

in the US inequality has increased substantially since the mid-1980s, there has been only a much more muted rise in the UK. As an instrument for the interest rate set by the Bank of England (BOE) we resort to the narratively identified shock series by Cloyne and Hürtgen (2016). They adopt C. Romer and D. Romer (2004)'s methodology to the UK to identify monetary policy surprises between 1975m1 and 2007m12.

Figures 2.6 and 2.7 show the estimated IRFs for industrial production and unemployment. Unconditionally, i.e. when we exclude the interaction terms with inequality in the regression equation, a contractionary monetary policy shock lowers industrial production and raises unemployment one to two years after its occurrence. While the dynamics are similar to those in the US, the magnitude of the responses is smaller in the UK, but well in line with those depicted in Cloyne and Hürtgen (2016).

Turning to the state dependence of monetary policy transmission, we obtain qualitatively similar results as for the US. Although at a few points during the first year the relationship is reversed, at most horizons changes in the interest rate tend to have more pronounced effects on real activity in times of high inequality. The absolute differences between high- and low-inequality responses, however, are a bit smaller in the UK. We plot the IRFs conditional on the minimum and the maximum observed level of inequality in our sample (a top 10% share of 45.6% and 61.0% respectively). An interest rate change that occurs in the so defined high-inequality regime lowers industrial production by up to 0.7% and raises unemployment by 0.16 percentage points, whereas under low inequality neither of the two variables shows pronounced deviations from zero at the considered horizons. In sum, an increase in the top 10% wealth share by one percentage point is associated with a 0.02 percentage points larger fall of industrial production over the first three years after the shock, and a 0.006 percentage points larger rise in unemployment.

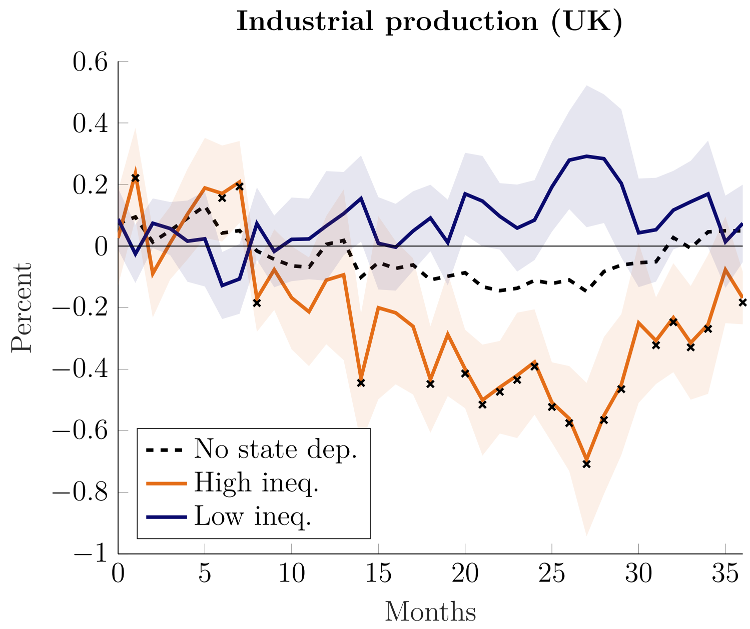


Figure 2.6: IRF of UK industrial production to a 25 bp change of the BOE policy rate.
Notes: Black dashed: No state dependence is imposed. Blue: Regime of low inequality (minimum of observed inequality). Red: Regime of high inequality (maximum). Confidence intervals are at the 90% level and based on Newey and West (1987) standard errors. Black crosses indicate horizons at which we reject $H_0 : \beta_h^+ = 0$ at a 90% confidence level.

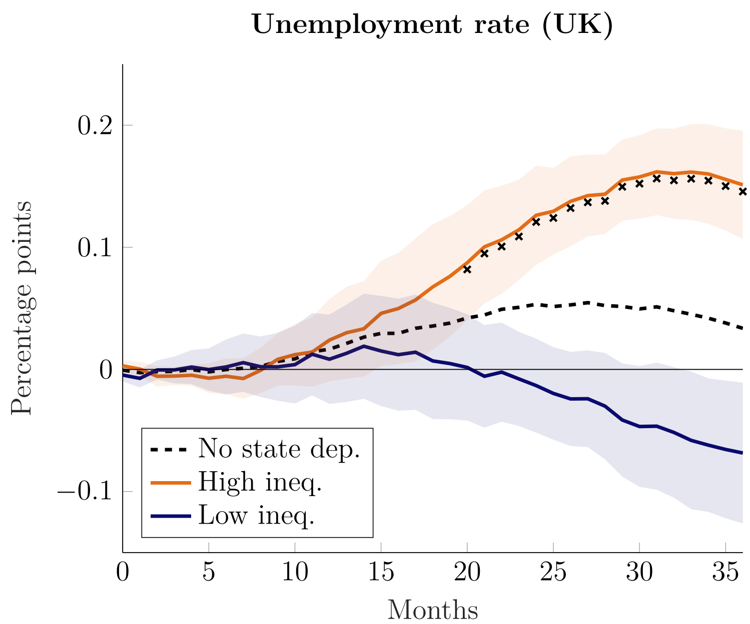


Figure 2.7: IRF of UK unemployment to a 25 bp change of the BOE policy rate.
Notes: Black dashed: No state dependence is imposed. Blue: Regime of low inequality (minimum of observed inequality). Red: Regime of high inequality (maximum). Confidence intervals are at the 90% level and based on Newey and West (1987) standard errors. Black crosses indicate horizons at which we reject $H_0 : \beta_h^+ = 0$ at a 90% confidence level.

3 Cross-Sectional Evidence from US States

While the advantages of using aggregate time series data for our analysis are the simplicity of the approach, the comparability to the literature and the high quality of the data, the main drawback lies in the difficulty to control for confounding factors. Any variable that evolved in a similar u-shape as the top 10% wealth share over the last decades could be responsible for our findings of state dependence. Therefore, we turn to a cross-sectional analysis in the next two sections, first of US states and then of Euro Area countries.

We start with the analysis based on US states. The main challenge we face here is that no measure of wealth inequality at the US state level is readily available. Before proceeding with our analysis we therefore construct our own measures, which we discuss next.

3.1 A measure of wealth inequality at the US state level

To measure wealth inequality at the state level we resort to publicly available data on top wealth holders derived from estate tax returns. The Federal estate tax was introduced in 1916 and is a tax on the transfer of wealth from the estate of a deceased person to its beneficiaries. A tax return must be filed by every deceased US citizen whose gross estate, valued on the date of death, equals or exceeds a certain exemption level, which has varied over time (see Jacobson et al. (2007) for a historical overview and detailed description of the Federal estate tax). Data on Federal estate tax returns have previously been used for research on wealth inequality (Kopczuk, 2015; Kopczuk and Saez, 2004; Saez and Zucman, 2016), but to the best of our knowledge not for a state-level analysis.

Based on estate tax returns, the Internal Revenue Service (IRS) periodically publishes estimates on the total wealth held by top wealth holders. Top wealth holders are defined as those currently alive US citizens for whom an estate tax return would be required upon death given the current exemption level regarding estate taxes. To arrive at the estimates of total wealth held by this group of individuals the IRS uses the estate multiplier technique. In short, every observed estate tax return is inflated by the inverse probability of that person dying in a given year. Therefore, old people (with relatively high probabilities of dying) receive low weights and young people receive high weights. This way, the IRS arrives at an estimate of the total wealth held by the wealthiest part of the population (see Jacobson et al. (2007) and Kopczuk and Saez (2004) for a detailed description of the methodology).

The IRS provides estimates on the total wealth held by top wealth holders for each US state in its Statistics of Income (SOI) Bulletins for the years 1976, 1982, and then every three years from 1986 onward. Given the total wealth held by top wealth holders as well the number of top wealth holders we construct the share of wealth held by the top 1% richest citizens for each state and year that is available. Appendix B lays out the details of this procedure. The key assumption that we make to arrive at our estimates is that at any given point in time and in each US state the right tail of the wealth distribution follows a Pareto distribution (Benhabib et al., 2015; Hubmer et al., 2021; Kopczuk and Saez, 2004). We construct top 1% shares here, instead of top 10% shares, since the exemption level for filing estate taxes has risen significantly over the recent decades and often times only covers a fraction of the top 1% of wealthiest citizens.[12] Extrapolating from this information to the wealth share of the top 10% is therefore difficult, and studies that use estate tax income for constructing inequality measures typically only report wealth shares of the top 1% or of even smaller fractions.

3.2 Empirical strategy

The self-constructed series of inequality measures on the US state level has one main drawback, namely that there are large gaps between the years for which we observe them. In the time between 1969 (when the Romer & Romer shock series begins) and 1986, we observe inequality only twice, and afterwards tri-annually. This precludes a time series analysis as we performed it on the national level where we observed inequality annually.

Instead, the approach we take in this section is to exploit cross-sectional variation between US states, similar to Carlino and DeFina (1998, 1999). For each state s we first estimate the IRF of a measure of economic activity in state s to a change in the interest rate by the FED. Similarly to the previous section we estimate IV local projections for horizons $h \in [0, 1, \ldots, H]$, but this time for each US state s:

$$y^s_{t+h} = \alpha^s_h + \beta^s_h \cdot i_t + \sum_{p=1}^{P} \Gamma^s_{h,p} \cdot X^s_{t-p} + u^s_{t+h} \tag{2.2}$$

The only difference to Section 2 is that we do not include interaction effects in the regression and that the outcome variable depends on the state s. The remaining control variables, the deterministic elements, as well as the number of included lags ($P = 2$) remain the same. Hence, X^s_t includes y^s_t, the aggregate unemployment rate, the federal funds rate, and the log of the two aggregate price indices.[13] As before, we instrument the federal funds rate i_t with the Romer & Romer shock series.

As a proxy for economic activity, we use two different outcome variables y^s_t at the state level, the log of real state personal income (as in Carlino and DeFina, 1998, 1999) and the state-level unemployment rate.[14] For the IRF of unemployment we use monthly data and consider a horizon of $H = 36$ months as in the previous section. Since state personal income is only observed quarterly, we estimate its IRF from quarterly data for horizons up to $H = 12$ quarters. We measure the impact of interest rate changes on economic activity as the average of the IRF over the first three years after the change in the interest rate, i.e.

[12] See Table 2.8 in the appendix for the cut-offs of wealth in between which the IRS provides its estimates regarding the top wealth holders.

[13] Including the logged sum of all remaining states' personal income as an additional control barely changes the results.

[14] We obtain all these series from FRED.

Table 2.1: Summary statistics for variables on the US state level (1969q1 to 2007q4).

	mean	sd	min	max
Avg. response personal income (in %)	-0.12	0.08	-0.27	0.18
Top 1% wealth share	0.24	0.04	0.14	0.36
Manuf. share	0.19	0.08	0.04	0.34
Share small firms	0.50	0.06	0.41	0.67
Share middle-aged	0.49	0.01	0.45	0.52
Share old	0.17	0.03	0.06	0.24
Observations	50			

we compute the cumulative response and then divide by $H+1$. In the following, we refer to this statistic as “impact measure”. Our sample runs from 1969q1 to 2007q4 in case of state personal income and from 1976m1 to 2007m12 for state unemployment, as observations on unemployment in the states are not available for earlier periods.

We regress our impact measure of monetary policy on the average top 1% wealth share in the sample period. In these regressions we can control for other variables that might affect the responses of economic activity to monetary policy and that could be correlated with wealth inequality. In particular, Carlino and DeFina (1998, 1999) find that a high share of income earned in the manufacturing sector of a state leads to larger responses to monetary policy shocks. They hypothesize that manufacturing is an interest-sensitive sector as purchases of housing, cars and other durable manufactured goods are relatively responsive to changes in the interest rate. Furthermore, in a subset of their regressions they find that the higher the percentage of small firms in a state the stronger the effects of monetary policy. Their proposed explanation is that small firms are more reliant on funding through banks and therefore more exposed to changes in the interest rate than large firms. Lastly, Leahy and Thapar (2019) find that US states with a high share of middle-aged people in the population (and to a lesser extent that of old-aged) react more strongly to monetary policy. The explanation they offer is that medium-aged people are relatively likely to be entrepreneurs and that therefore private firm investment becomes more responsive in states where the share of this demographic group is large.

We investigate these proposed channels by including the share of state income that is earned in the manufacturing industries, the share of employees who work in firms with less than 250 workers, and the share of the population aged 35–65 and the share older than 65 years in a subset of our regressions.[15] For all variables we take the average over the respective sample period. Table 2.1 displays summary statistics of the variables we use.

[15]For the manufacturing share we divide earnings in the manufacturing industries by total earnings in a state, both provided by the Bureau of Economic Analysis at quarterly frequency. Employment shares are computed on the basis of the Business Dynamics Statistics provided by the U.S. Census Bureau on an annual basis from 1977 onward. Demographic groups are computed using annual data from the U.S. Census Bureau starting in 1970.

Table 2.2: Regression results for the cross-section of US states (state personal income).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Top 1% wealth share	-0.4	-0.5**	-0.4	-0.2	-1.3***	-1.4***	-1.3***	-1.1**
	(0.3)	(0.2)	(0.2)	(0.2)	(0.4)	(0.4)	(0.5)	(0.5)
Manuf. share		-0.7***	-0.6***	-0.4**		-1.1***	-1.1**	-0.7
		(0.1)	(0.2)	(0.2)		(0.4)	(0.5)	(0.5)
Share small firms			0.2	0.4*			0.2	0.9*
			(0.2)	(0.2)			(0.5)	(0.5)
Share middle-aged				-1.8***				-6.0***
				(0.6)				(1.4)
Share old				-0.9*				-1.8**
				(0.5)				(0.9)
Observations	50	50	50	50	50	50	50	50
Sample	'69–'07	'69–'07	'69–'07	'69–'07	'84–'07	'84–'07	'84–'07	'84–'07

Notes: Dependent variable is the average IRF of log real state personal income to a 25 bp increase in the fed. funds rate over a horizon of three years. Robust standard errors are reported in parentheses, a * indicates $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

3.3 Results

The first four columns of Table 2.2 show the results of our regressions when using the average IRF of state personal income as the independent variable. In the first column we only use wealth inequality as an explanatory variable, and we subsequently add the remaining explanatory variables in columns 2 to 4. In line with the results in the previous section a negative relationship between the average response and wealth inequality emerges. The point estimate in column 1 indicates that an increase of the top 1% wealth share by one standard deviation (0.044) decreases the state's average impulse response to a contractionary monetary policy shock of 25 basis points by $0.044 \cdot 0.4 = 0.02$ percentage points over the first three years after the shock. In line with Carlino and DeFina (1998, 1999) we find a negative impact of a high manufacturing share, while in contrast to them we find that a high share of employment in small firms mutes the output response. A large fraction of middle-aged (and to a lesser extent that of old-aged) leads to stronger responses to interest rate changes, as in Leahy and Thapar (2019).

The first four columns of Table 2.3 repeat the analysis when using the average IRF of state unemployment as our impact measure. In line with our previous findings, we estimate a positive coefficient on the top 1% wealth share, implying a stronger rise in unemployment in states where wealth is more unequally distributed. Column 1 implies that the average rise in unemployment becomes $0.044 \cdot 0.1 = 0.004$ percentage points larger when a state's top 1% wealth share increases by one standard deviation.[16]

While the point estimates of the coefficient on the top 1% wealth share in the first four

[16]The standard deviation of the average unemployment response is 0.04 percentage points, its mean 0.08 percentage points. All summary statistics of the sample starting in 1976q1 are shown in Table 2.9 in the appendix.

Table 2.3: Regression results for the cross-section of US states (unemployment).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Top 1% wealth share	0.1	0.2	0.08	0.07	0.5**	0.6***	0.5*	0.4*
	(0.1)	(0.1)	(0.1)	(0.1)	(0.2)	(0.2)	(0.3)	(0.2)
Manuf. share		0.4***	0.3***	0.3***		0.4*	0.3	0.2
		(0.06)	(0.08)	(0.09)		(0.2)	(0.2)	(0.2)
Share small firms			-0.2	-0.2			-0.2	-0.4*
			(0.1)	(0.1)			(0.3)	(0.2)
Share middle-aged				0.05				2.1***
				(0.3)				(0.7)
Share old				0.006				0.3
				(0.2)				(0.4)
Observations	50	50	50	50	50	50	50	50
Sample	’76–’07	’76–’07	’76–’07	’76–’07	’84–’07	’84–’07	’84–’07	’84–’07

Notes: Dependent variable is the average IRF of the state unemployment rate to a 25 bp increase in the fed. funds rate over a horizon of three years. Robust standard errors are reported in parentheses, a * indicates $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

columns of Tables 2.2 and 2.3 all have the expected sign, the results for wealth inequality are statistically significant only in a subset of cases. One possible explanation for this is that time-varying inequality on the state level confounds the results. In particular, as was the case for the US as a whole, our measures of wealth inequality for each state are not constant but vary over time. By taking the average over time for each state we effectively assume a constant degree of wealth inequality over the entire sample period.

We therefore repeat our analysis, this time restricting our sample to start in 1984q1. We choose this starting point for two reasons. First, as highlighted by Coibion (2012) and as explained in Section 2, the Federal Reserve targeted non-borrowed reserves instead of the federal funds rate in the early 1980s, which introduces some noise into the estimation of IRFs to changes in the interest rate. Owyang and Wall (2009) furthermore find that the regional effects of monetary policy in the US in the Volcker-Greenspan era post 1983 differed significantly from its effects in earlier episodes. Second, estimates on the top wealth holders are consistently published by the IRS on a tri-annual basis only from 1986 onward. Hence, when using the full sample there exists a long period (1969–1985) that is used for estimating the state-level IRFs but during which we observe wealth inequality only twice. When using the sample starting in 1984 we instead observe inequality at a constant frequency of three years.

The last four columns of Tables 2.2 and 2.3 display the results for the sample from 1984q1 to 2007q4. In all specifications we find a statistically significant relationship between the top 1% share and the state's response to monetary policy. Moreover, the point estimates are larger than those obtained using the full sample. We find that an increase in the top 1% wealth share by 4.4 percentage points exacerbates the contraction of real state personal income by $0.044 \cdot 1.3 = 0.06$ percentage points on average over a horizon of three years. The

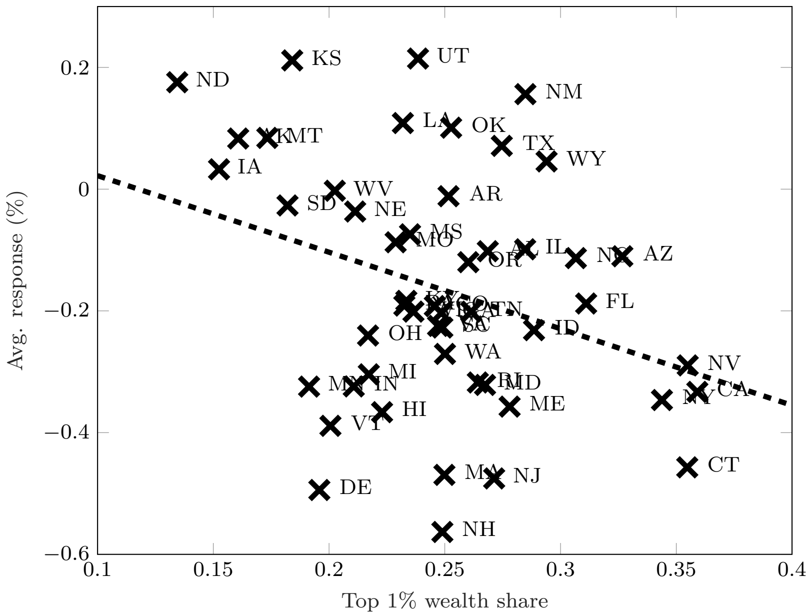


Figure 2.8: Top 1% shares and average IRFs of personal income.

Notes: Top 1% shares are averaged over the sample. Average IRFs are taken over the first three years after the shock. Sample period is 1984q1–2007q4.

same increase in the top 1% share now on average implies that the increase in unemployment becomes $0.044 \cdot 0.5 = 0.022$ percentage points larger. Figure 2.8 shows a scatter plot of the states' average top 1% shares and their cumulative IRFs for the sample starting in 1984q1.

3.4 Robustness

First, we use the State Coincident Index calculated by the Federal Reserve Bank of Philadelphia as our measure of activity on the state level instead of state personal income or unemployment. The Coincident Index is available at monthly frequency from 1979m1 onward and is computed on the basis of four variables, in particular nonfarm payroll employment, average hours worked in manufacturing by production workers, the unemployment rate, and wage and salary disbursements deflated by the consumer price index (U.S. city average). The results, using the average IRF over the first $H = 36$ months after the shock as the impact measure, are shown in Table 2.10 in Appendix B.2. They are very similar to those obtained when using state personal income as the measure of activity (Table 2.2).

Second, we measure the impact of monetary policy at the state level not with the average but the peak response of state personal income. The results of this are shown in Table 2.11 in Appendix B.2. While over the entire sample we obtain estimates that are not statistically different from zero, when focusing on the sample starting in 1984 we again find that responses are more pronounced in states with high inequality.

Lastly, we exploit the length of the time series and construct a two-period panel of state-level observations. In particular, we split the sample into two subsamples of equal length, 1969q1–1988q2 and 1988q3–2007q4, and for every state estimate both the average top 1% wealth share as well as IRFs to monetary policy shocks separately for the two subsamples.

We use the resulting panel of average top 1% wealth share and impact measure of monetary policy to estimate a model with state fixed effects. This allows us to control for all time-invariant characteristics of a state that might affect its response to interest rate changes. Table 2.12 in the appendix displays the resulting estimates when using the average IRF of state personal income as our impact measure. Even when controlling for state fixed effects a negative point estimate on the top 1% share emerges, though it is not statistically significant (columns 3–6). The first two columns also display the estimated coefficient of the top 1% share for each of the two subsamples. Both are negative and statistically significant. While this finding is not surprising for the second subsample (1988–2007) as it almost coincides with the subsample used in the last three columns of Table 2.2 (1984–2007), it is encouraging to see that also before 1988 we find evidence for a state dependence of monetary policy transmission on wealth inequality in the US states.

4 Cross-Sectional Evidence from the Euro Area

Next, we analyze the correlation between wealth inequality and the real effects of monetary policy in the cross-section of Euro Area countries. Our empirical approach is similar to the one for US states in Section 3 and to the analysis in Almgren et al. (2020). The additional advantage of using Euro Area data compared to the US state-level analysis is the availability of a micro data set, the HFCS, that has detailed information on households' asset holdings and is representative at the country level. Hence we can control for a host of potential confounding factors that may be correlated with wealth inequality, such as the share of hand-to-mouth households and the rate of home ownership. As for the US states, we have to estimate the effects of monetary policy on economic activity as well as wealth inequality for every Euro Area country.

4.1 The effects of monetary policy

We first estimate the real effects of monetary policy for all nineteen Euro Area countries. For every country c, we estimate IV local projections as before and use monthly Eurostat data from 1999m1, when the Euro was introduced, to 2020m1. We estimate for $h \in [0, 1, \ldots, H]$

$$y^c_{t+h} = \alpha^c_h + \beta^c_h \cdot i_t + \sum_{p=1}^{P} \Gamma^c_{h,p} \cdot X^c_{t-p} + u^c_{t+h} \,. \tag{2.3}$$

As before, y^c_{t+h} is the outcome variable of interest in country c at time $t+h$, i_t is the nominal interest rate and X^c_t is a vector of controls. X^c_t contains country c's unemployment rate, the log of its GDP, the log of its CPI, the identified monetary policy shocks described below, and the nominal interest rate for which we take the Euro Over Night Index Average (EONIA) rate.[17] As before, we estimate IRFs up to a horizon of $H = 36$. We follow Almgren et al. (2020) in the choice of a lag length of $P = 3$ and α_h^c being a (country- and horizon-specific) constant. We resort to monthly frequency, however GDP for the Euro Area countries is only available at a quarterly frequency. To obtain a monthly series, we follow Almgren et al. (2020), who employ the Chow and Lin (1971) procedure using data on unemployment, industrial production and retail trade to interpolate quarterly GDP.

We instrument the nominal interest rate using shocks identified from high-frequency movements in Overnight Indexed Swaps (OIS). In particular, we construct the shocks from changes in the 3-months OIS rates around ECB monetary policy announcements, which are provided by Altavilla et al. (2019). Since ECB monetary policy decision are announced at 13:45 followed by a press conference that ends at 15:30, Altavilla et al. (2019) compute the difference in the median price of the OIS in the time between 15:40 to 15:50 and the median price between 13:25 to 13:35. Under the identifying assumption that policy decisions are the only relevant factor driving interest rates in this short time window, these changes constitute exogenous variations in nominal interest rates and are therefore valid instruments.

This procedure yields a daily series of monetary policy shocks that we have to aggregate to monthly frequency. To do this, we follow the procedure employed in Meier and Reinelt (2020) and Ottonello and Winberry (2020). The idea is to account for the fact that for a shock that takes place late in a month little time is left to affect the economy. For example, a shock on the last day of a month is much closer in time to a shock that takes time on the first day of the next month than it is to a shock on the first day of the same month and we would expect to see this reflected in how the economy responds to the shock. To take this into account, we attribute each shock in part to the next month depending on how many days are left in the month after the day of the shock. In particular, the value of the monthly series in month t, denoted ϵ_t, is constructed according to

$$\epsilon_t = \sum_{\tau \in D(t)} \phi(\tau, t)\tilde{\epsilon}_\tau + \sum_{\tau \in D(t-1)} (1 - \phi(\tau, t-1))\tilde{\epsilon}_\tau,$$

where $\tilde{\epsilon}_\tau$ is the value of the daily series on day τ, $D(t)$ is the set of days in month t and $\phi(\tau, t)$ is the share of days of the month after day τ. In words, a shock that occurs on the first day of a month is fully attributed to that month while for a shock that occurs in the middle of the month half is attributed to the current month and half to the next month.

Not all current Euro Area countries have been members of the Euro Area since its inception in 1999 such that some were not directly subject to ECB monetary policy in the earlier years of our sample.[18] Nevertheless, for our baseline specification we estimate impulse responses to ECB interest rate changes for all countries based on the whole time period. We

[17] As Almgren et al. (2020) we include the shocks as controls to alleviate the problem of serial correlation in the series of shocks.

[18] Slovenia joined in 2007, Cyprus and Malta in 2008, Slovakia in 2009, Estonia in 2011, Latvia in 2014 and Lithuania in 2015.

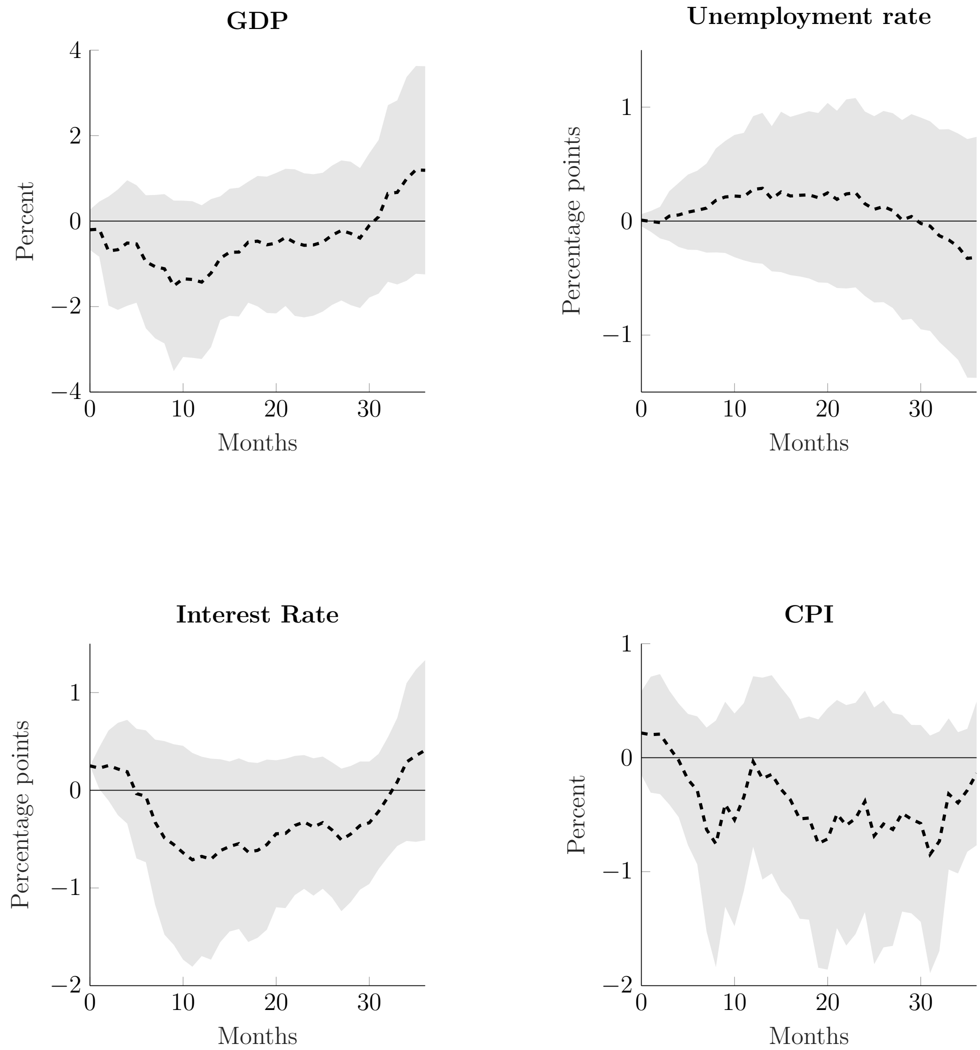


Figure 2.9: IRFs to a 25 bp increase in the nominal rate for the Euro Area as a whole.
Notes: Shaded areas indicate 90% confidence intervals based on Newey and West (1987) standard errors.

think this approach is reasonable because even those countries who joined later had already pegged their exchange rates to the Euro much earlier and were thus strongly affected by the ECB's monetary policy. As a robustness check we also estimate IRFs only based on those time periods in which a country was a member of the Euro Area. We also consider only the sub-sample of the original eleven Euro Area countries (EA11).

We first estimate (2.3) for the Euro Area as a whole. Figure 2.9 shows the estimated IRFs to a 25 basis points change in the nominal interest rate. In response to the interest rate hike, both GDP and unemployment respond as expected. GDP falls, reaching a trough at about -1.5% after one year. Unemployment responds more slowly initially but then peaks at about 0.3 percentage points. The response of GDP is larger than that of industrial production we estimated for the US (Figure 2.2, black dashed line) but it is similar to the results in Almgren et al. (2020).[19]

[19] Almgren et al. (2020) display IRFs of GDP to an intererst rate reduction by one standard deviation of the shock series. The standard deviation of our constructed shock series is 3 basis points, or 0.0003.

4.2 A measure of wealth inequality at the country level

We next turn to the measurement of wealth inequality at the country level. We estimate the top 10% wealth share country by country from the HFCS. The HFCS is a survey of households finances in EU countries similar to the Survey of Consumer Finances in the US. It has been conducted in three waves between 2010 and 2017 by the individual member states and is representative on the national level. Importantly, the core questions of the survey used for our analysis are identical across countries which allows for an easy comparison between countries. To obtain an estimate of the top 10% wealth share for every EA country, we define net worth for every household as the difference between total assets excluding public and occupational pensions plans and total outstanding household liabilities.[20] Total assets is the sum of total real assets, which consist of real estate, vehicles, self-employed businesses, and other valuables, and total financial assets, consisting of deposits, mutual funds, bonds, non-self-employed businesses, shares, managed accounts, money owed to the household, life-insurance and other pensions, and other assets. Total liabilities consist of mortgage debt and other debt such as credit card debt and consumer loans.[21]

We compute the top 10% wealth share based on this definition of wealth for every country and each HFCS wave and then average over the three HFCS waves to obtain a single measure of wealth inequality for every country. The resulting measure is depicted in Figure 2.10. In the average Euro Area country, the richest ten percent of households hold about 50% of total net worth as indicated by the vertical line. However, there is large variation across countries. The top 10% wealth share is largest in Germany with an average of 59% between 2010 and 2017. It is lowest in Slovakia and Greece at 36% and 41% respectively.

4.3 Results

Table 2.4 reports summary statistics for both the average GDP and unemployment responses to a 25 bp monetary policy shock as well as the top 10% wealth shares for all 19 Euro Area countries. The means of the average GDP and unemployment responses are in line with the IRFs for the Euro Area as a whole (Figure 2.9). Compared to the US states, the responses in the Euro Area countries display a much higher standard deviation. This might owe in part to the much shorter sample period over which we estimate the IRFs here.

We now ask whether the observed differences in wealth inequality can explain the differential responses to monetary policy. The first column in Table 2.5 shows the result of a linear regression of the average GDP response on the top 10% wealth share. As for the US states, we estimate a negative relationship between inequality and the output response to an increase in the interest rate. The point estimate indicates that an increase in the top

[20]For the construction of our measure of wealth inequality, as well as for all other variables we use in our analysis below that are derived from the HFCS, we only consider households whose head is aged 20–75.

[21]This definition corresponds to variable DN3001 (Net wealth excl. public and occupational pensions) in the HFCS.

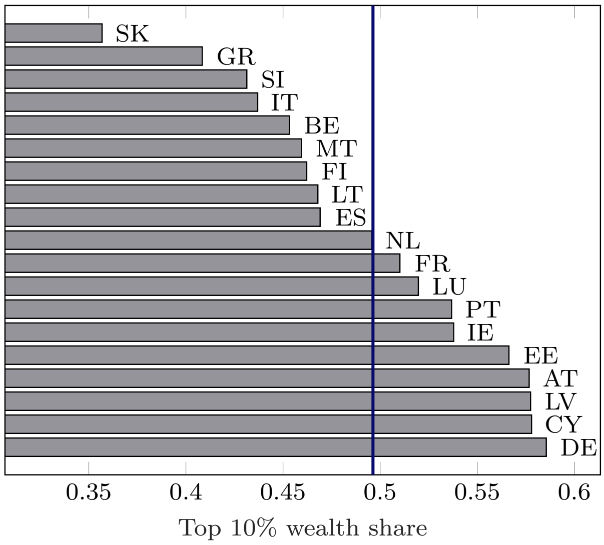


Figure 2.10: Top 10% wealth share in Euro Area countries.

Notes: Top 10% wealth shares, averaged across all waves of the HFCS for which data on net worth exists in a given country. Solid line: Average across countries.

Table 2.4: Summary statistics for Euro Area countries (1999m1 to 2020m1).

	mean	sd	min	max
Avg. response GDP (in %)	-0.47	0.86	-2.26	1.37
Avg. response unemployment (in p.p.)	0.05	0.45	-0.85	0.93
Top 10% wealth share	0.50	0.07	0.36	0.59
Observations	19			

Table 2.5: Regression of avg. responses to monetary policy shocks on wealth inequality.

	(1) GDP	(2) Unemp.	(3) U EA11	(4) U when in EA	(5) U peak	(6) GDP	(7) Unemp.
Top 10% share	-4.6*	4.1***	3.7	4.9**	4.7**		
	(2.5)	(1.0)	(2.1)	(2.0)	(1.8)		
Top 1% share						-6.1	3.9**
						(3.7)	(1.6)
Observations	19	19	11	19	19	19	19

Notes: Dependent variable is the average IRF of GDP (in %) or unemployment (in p.p.) to a 25 bp increase of the interest rate over a horizon of three years. Robust standard errors are reported in parentheses, a * indicates $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

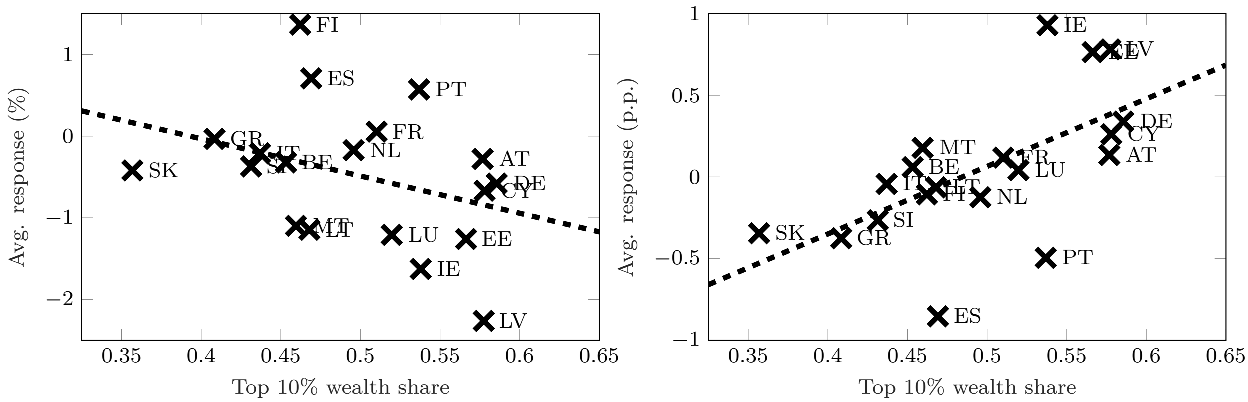


Figure 2.11: Top 10% wealth share and avg. response of GDP (left) and unemployment (right).

10% wealth share by one standard deviation is associated with a $0.07 \cdot 4.6 = 0.32$ percentage points stronger contraction in GDP in response to a 25 basis points shock.

In the second column we repeat the analysis, this time using the average response of unemployment as the impact measure. The estimated effect is positive, as would be expected from our previous results, and statistically significant at the 1% level. The point estimate implies that an increase in the top 10% wealth share of one standard deviation is associated with an increase in the unemployment response to a 25 basis points rate hike by $0.07 \cdot 4.1 = 0.29$ percentage points. Given that the standard deviation of average unemployment responses across countries is 0.45 percentage points, this is a substantial effect. We return to a discussion of the remaining columns of Table 2.5 below. Figure 2.11 shows the correlation between the degree of wealth inequality in Euro Area countries and their average response of GDP (left panel) and unemployment (right panel).

Additional controls There are several alternative explanations for the observed differences in the response to monetary policy changes and the availability of comparable country-level data in the Euro Area allows us to investigate some of them by including additional

Table 2.6: Regressions of impact measures on wealth inequality and controls

	(1) GDP	(2) UR	(3) GDP	(4) UR	(5) GDP	(6) UR	(7) GDP	(8) UR	(9) GDP	(10) UR
Top 10% share	-4.1*	4.0***	-5.4	4.3***	-7.1**	4.4***	-3.6	4.2***	-6.9**	5.0***
	(2.0)	(0.9)	(3.4)	(1.2)	(3.0)	(1.3)	(2.1)	(1.0)	(2.5)	(1.0)
Htm share	-3.0**	0.8*								
	(1.3)	(0.5)								
Empl. protection			0.3	-0.3**						
			(0.3)	(0.1)						
Home ownership					-3.0**	0.3				
					(1.3)	(0.7)				
Share adj. mortg.							-0.7	-0.1		
							(0.6)	(0.3)		
Debt-to-GDP									1.6**	-0.6*
									(0.7)	(0.3)
Observations	19	19	16	16	19	19	18	18	19	19

Notes: Dependent variable is the average IRF of GDP (columns with "GDP") or unemployment ("UR") to a 25 bp increase of the interest rate over a horizon of three years. Robust standard errors are reported in parentheses, a * indicates $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

controls.[22] Since our sample is substantially smaller than on the US state level, we only include one control at a time.

First, we add the share of hand-to-mouth households, i.e. of those households who hold only very few liquid assets, as a control. Kaplan, Moll, et al. (2018) argue that a higher share of liquidity constrained households raises the aggregate marginal propensity to consume and thereby strengthens the effects of monetary policy. Almgren et al. (2020) find a positive relationship between the hand-to-mouth share and the impact of monetary policy using a very similar strategy to the one we take in this section. Using the same definition of hand-to-mouth households as them, the first two columns in Table 2.6 confirm their findings: both GDP and the unemployment rate respond more strongly to interest rate changes in countries with a higher share of hand-to-mouth households. However, the estimated coefficients on wealth inequality do not change much relative to the baseline when we control for the hand-to-mouth share, and they remain statistically significant.

Next, labor market institutions may be important for the response to monetary policy, especially when we use the unemployment rate as a proxy for economic activity. We therefore add an index of employment protection strictness constructed by the OECD (2020) as an additional control.[23] The index measures the strictness of regulation on collective dismissals, and high values indicate a stronger protection of the workforce. As shown in columns 3 and 4 we only find a statistically significant (dampening) effect on the unemployment rate and

[22]We obtain data on household debt to GDP, the share of hours worked in the manufacturing sector, and the share of middle-aged households from Eurostat. We construct the remaining controls from the HFCS, unless stated otherwise.

[23]The index is not available for Cyprus, Lithuania and Malta.

Table 2.7: Regressions of impact measures on wealth inequality and controls (cont.)

	(1) GDP	(2) UR	(3) GDP	(4) UR	(5) GDP	(6) UR	(7) GDP	(8) UR
Top 10% share	-4.7*	4.3***	-4.6	4.1***	-6.0**	4.1***	-5.1**	4.2***
	(2.6)	(1.1)	(2.7)	(1.1)	(2.5)	(1.1)	(2.1)	(1.1)
Manuf. share	-1.6	1.8						
	(2.8)	(1.4)						
Share small firms			-3.1	0.2				
			(2.2)	(0.8)				
Share middle-aged					-10.2**	-0.5		
					(4.6)	(2.0)		
Stock mark. participation							4.1**	-0.3
							(1.7)	(0.6)
Observations	19	19	19	19	19	19	19	19

Notes: Dependent variable is the average IRF of GDP (columns with "GDP") or unemployment ("UR") to a 25 bp increase of the interest rate over a horizon of three years. Robust standard errors are reported in parentheses, a * indicates $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

adding the index as a control does not change the conclusion about the role of wealth inequality.

Beraja et al. (2019) and Wong (2021) suggest that the rate of home ownership and the share of adjustable rate mortgages is important for the transmission of monetary policy. Home owners are likely to have mortgages and are therefore particularly affected by interest rate changes. This effect could be even stronger if many mortgages have adjustable rates since monetary policy has a direct effect on existing mortgages in this case. We thus control for the home ownership rate and the share of households with adjustable rate mortgages in columns 5 to 8.[24] The results we obtain are mixed. If anything, a higher share of home owners leads to stronger effects of monetary policy as expected, but we do not find a significant role for the share of adjustable rate mortgages.

For the US, Alpanda and Zubairy (2019) find evidence that higher levels of household debt are associated with smaller effects of monetary policy shocks. Columns 9 and 10 show that this relationship also holds in the cross-section of Euro Area countries, which corroborates the findings in Alpanda and Zubairy (2019). The point estimates for the effect of wealth inequality become even larger when controlling for household debt.

Next, as described in the previous section, regions with a large manufacturing industry might respond more strongly to monetary policy. We do not find a statistically significant effect of the percentage of total hours worked in the manufacturing sector, though the point estimates have the expected sign (Table 2.7, columns 1 and 2). In contrast to the results on the US state level, we find that a large share of employment in small firms is associated with larger effects of monetary policy, though the relationship is not statistically significant (columns 3 and 4).

[24] Data on mortgage types are not available for Finland.

As explained before, Leahy and Thapar (2019) find demographics to be important for the transmission of monetary policy in the US, so we control for the share of the population aged 35–65 in columns 5 and 6. We find a statistically significant amplifying effect of the share of middle-aged households on the response of GDP also in the Euro Area, but there is no significant effect on the unemployment response. Lastly, we control for stock market participation, measured as the share of households that either hold stocks directly or via mutual funds. Contrary to the results in Melcangi and Sterk (2020), we find that, if anything, higher stock market participation weakens the effects of monetary policy on real activity.

In sum, many of the factors that were found to affect the strength of the monetary transmission mechanism turn out to be important also in our data. Wealth inequality, however, appears to have an effect on monetary policy transmission beyond its working through either of these factors.

4.4 Robustness

We conduct a number of robustness checks and list the results in columns 3 to 7 of Table 2.5. First, in column 3 we restrict the sample to the original eleven member states. We estimate a similar effect on the unemployment response as in the baseline but since the sample only consists of eleven countries in this case, it is not statistically significant. The fourth column corresponds to the case where responses to monetary policy are estimated using only those country-month observations where a given country was a member of the Euro Area. Put differently, for those countries that joined later we do not use the months before it adopted the Euro. In this case we find an even stronger effect of wealth inequality on the effectiveness of monetary policy but the estimate also has a larger standard error.

Next, in column 5 we look at the peak response of unemployment instead of its average response. The point estimate in this case is similar to the estimate for the average response. Lastly, we regress our impact measures not on the top 10% but on the top 1% wealth share (columns 6 and 7). Though the estimated coefficients have larger standard errors in this case our conclusion about the correlation between wealth inequality and the strength of monetary policy transmission remains intact.

5 Conclusion

This paper was motivated by two strands of literature that have received a lot of attention over the past decade. The first one documents significant variation across countries and time in the degree of wealth inequality in the population. The second emphasizes the role of household heterogeneity for short-run phenomena such as the transmission of monetary policy. We studied empirically how the strength of monetary policy transmission depends on the degree of wealth inequality, and we found that the effects of interest rate changes are state-dependent. More unequal distributions of household wealth are associated with stronger effects of monetary policy on real variables, such as GDP, industrial production and unemployment. This relationship holds in all three contexts we considered, on the aggregate level in the US and the UK, in the cross-section of US states and in the cross-section of Euro Area countries.

Our empirical analysis implies that the distribution of wealth matters and that the effects of interest rate changes are not invariant to changes in the level of inequality. This provides a rationale for continuing on the path of using Heterogeneous Agent New Keynesian (HANK) models for analyzing monetary policy. Our study can help inform the future building of such models. Analyzing further which underlying forces are causing our empirical results as well as designing HANK models that are in line with them appear like interesting avenues for future research.

Policymakers in central banks take an increasing interest in the interaction of monetary policy and household heterogeneity (BIS, 2021; Dossche et al., 2021). As mentioned in the beginning, our findings are relevant for their current discussions. Based on our results, the negative output and employment effects of raising interest rates are larger today when wealth inequality is at historically high levels than they were on average in the past. Concerns that the ongoing Covid-19 pandemic might further deepen the inequalities in advanced economies further add to the relevance of our findings (Chetty et al., 2020; Kartashova and Zhou, 2021).

Appendices

A Aggregate Time Series

A.1 Identifying assumptions

For the estimates of β_h to be unbiased we require our instrument to be exogenous. Formally, denoting by RR_t the Romer & Romer shock at time t we obtain

$$Cov(RR_t, u_t) = 0$$

by assuming $\mathbb{E}[RR_t|u_t] = \mathbb{E}[RR_t]$ and the fact that we have assumed $\mathbb{E}[u_t] = 0$. That contemporaneous outcome variables do not contain information about the Romer & Romer shocks, i.e. $\mathbb{E}[RR_t|u_t] = \mathbb{E}[RR_t]$, is a reasonable and common assumption in the literature and follows from the identification strategy of C. Romer and D. Romer (2004).

For the estimates of β_h^+ to be unbiased, we require (given that $\mathbb{E}[u_t] = 0$):

$$\begin{aligned} Cov(RR_t \cdot \text{ineq}_t, u_t) &= \mathbb{E}[RR_t \cdot \text{ineq}_t \cdot u_t] \\ &= \mathbb{E}\left[u_t \cdot \mathbb{E}\left[RR_t \cdot \mathbb{E}\left[\text{ineq}_t|RR_t\right]|u_t\right]\right] = 0 \ . \end{aligned}$$

This would be fulfilled if we had that $\mathbb{E}[\text{ineq}_t|RR_t] = \mathbb{E}[\text{ineq}_t]$. Given the results in Coibion et al. (2017), however, one might suspect a contemporaneous relationship between monetary policy and wealth inequality. Therefore, we lag inequality by one year (twelve months) when using it as an instrument and assume $\mathbb{E}[\text{ineq}_{t-12}|RR_t] = \mathbb{E}[\text{ineq}_{t-12}]$. Since past inequality cannot react to contemporaneous monetary policy shocks the only assumption we make is therefore that the FOMC does not react to wealth inequality in the past. This seems like a reasonable assumption for two reasons. First and most importantly, the Federal Reserve's mandate does not include wealth inequality and its policy should therefore not be expected to depend on wealth inequality directly. Second, inequality has only recently received attention by central bankers. It was not so much part of policy discussions in the time before 2007 on which our estimates are based (see for example BIS, 2021).

A.2 Additional graphs and figures

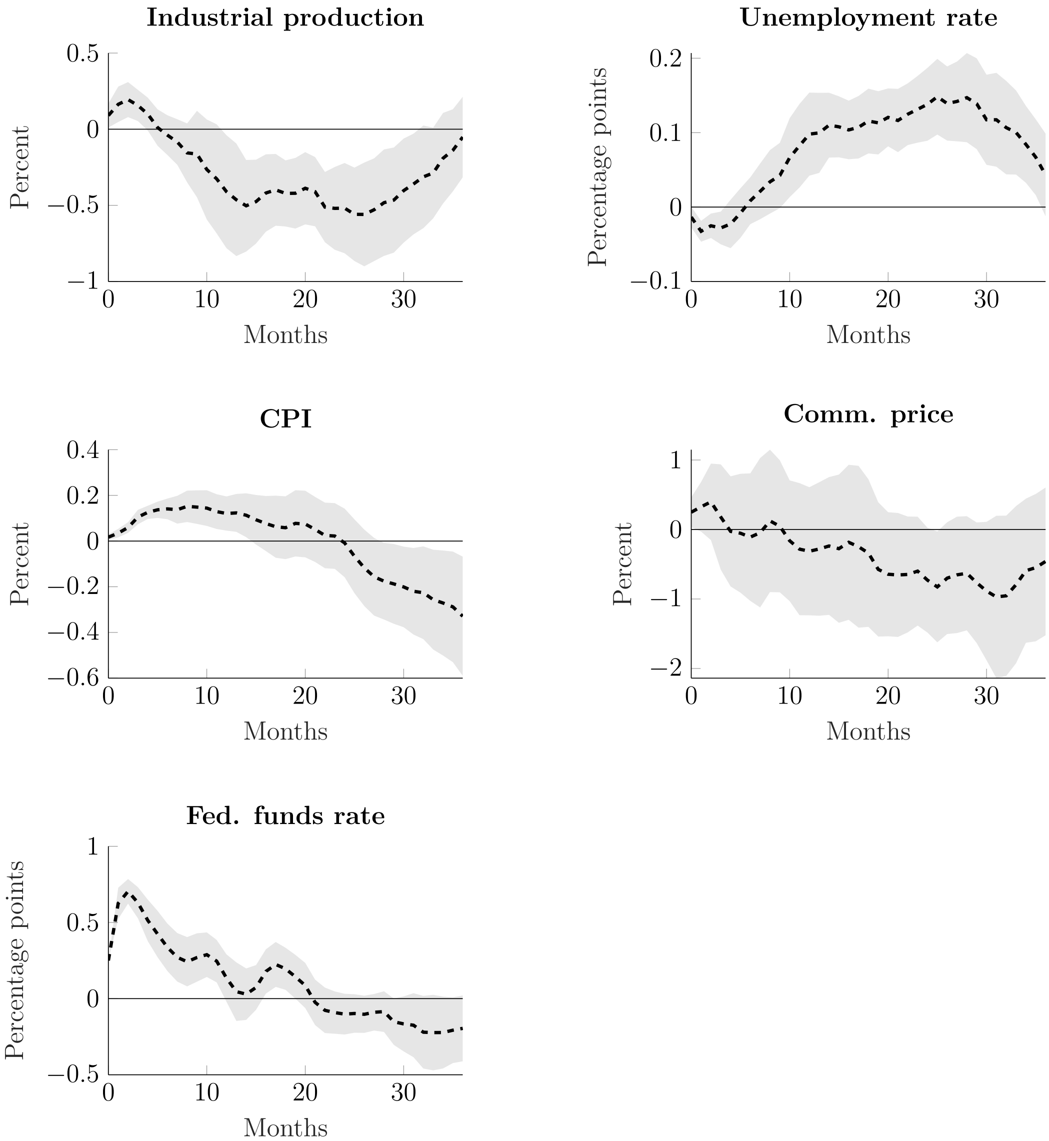


Figure 2.12: Unconditional IRFs to a 25 basis points change of the fed. funds rate.
Notes: Depicted are 90% confidence intervals based on Newey and West (1987) standard errors.

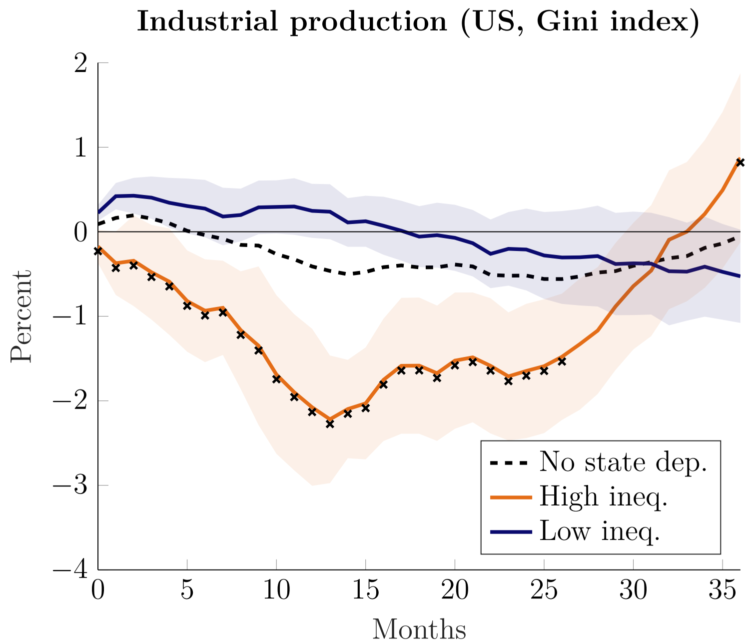


Figure 2.13: Impulse response to a 25 basis points change of the fed. funds rate.
Notes: ineq_t is measured by the Gini index. Black dashed: No state dependence is imposed. Blue: Regime of low inequality (first quartile of observed inequality). Red: Regime of high inequality (third quartile). Confidence intervals are at the 90% level and based on Newey and West (1987) standard errors. Black crosses indicate horizons at which we reject $H_0 : \beta_h^+ = 0$ at a 90% confidence level.

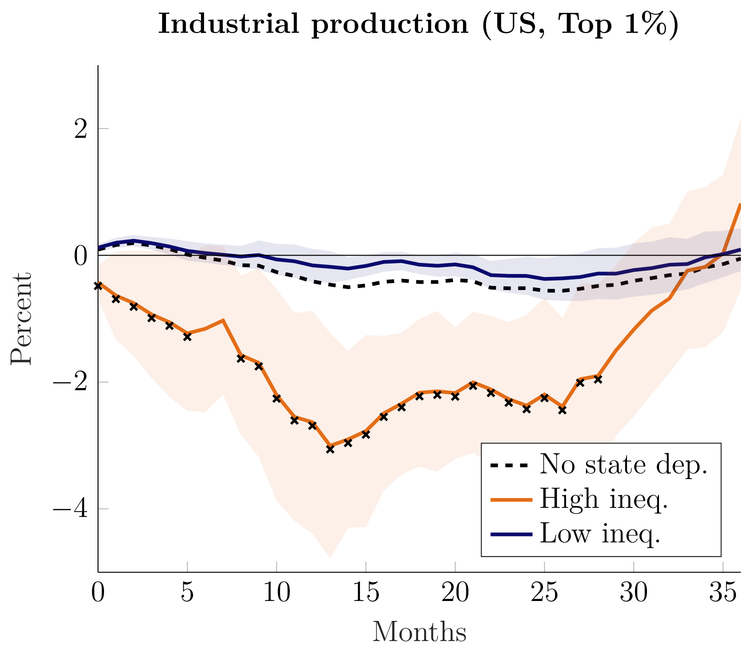


Figure 2.14: Impulse response to a 25 basis points change of the fed. funds rate.
Notes: ineq_t is measured by the top 1% wealth share. Black dashed: No state dependence is imposed. Blue: Regime of low inequality (first quartile of observed inequality). Red: Regime of high inequality (third quartile). Confidence intervals are at the 90% level and based on Newey and West (1987) standard errors. Black crosses indicate horizons at which we reject $H_0 : \beta_h^+ = 0$ at a 90% confidence level.

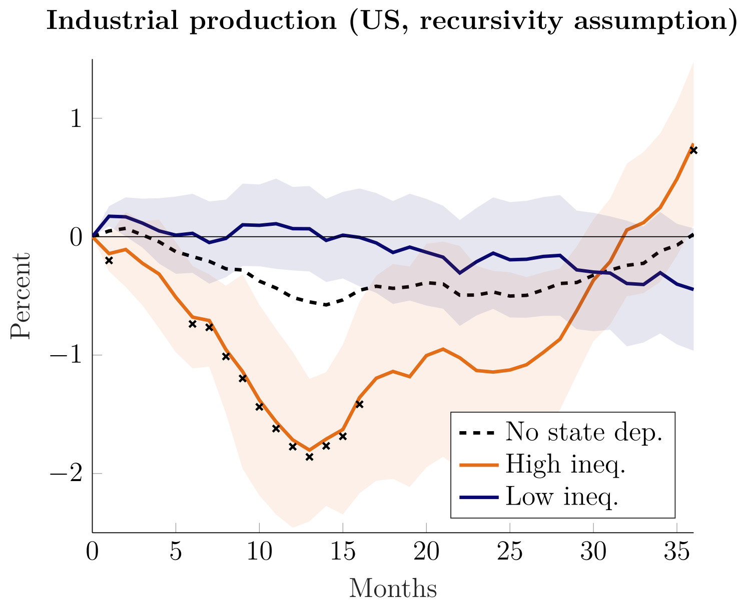


Figure 2.15: Impulse response to a 25 basis points change of the fed. funds rate.
Notes: Recursivity assumption is imposed. Black dashed: No state dependence is imposed. Blue: Regime of low inequality (first quartile of observed inequality). Red: Regime of high inequality (third quartile). Confidence intervals are at the 90% level and based on Newey and West (1987) standard errors. Black crosses indicate horizons at which we reject $H_0 : \beta_h^+ = 0$ at a 90% confidence level.

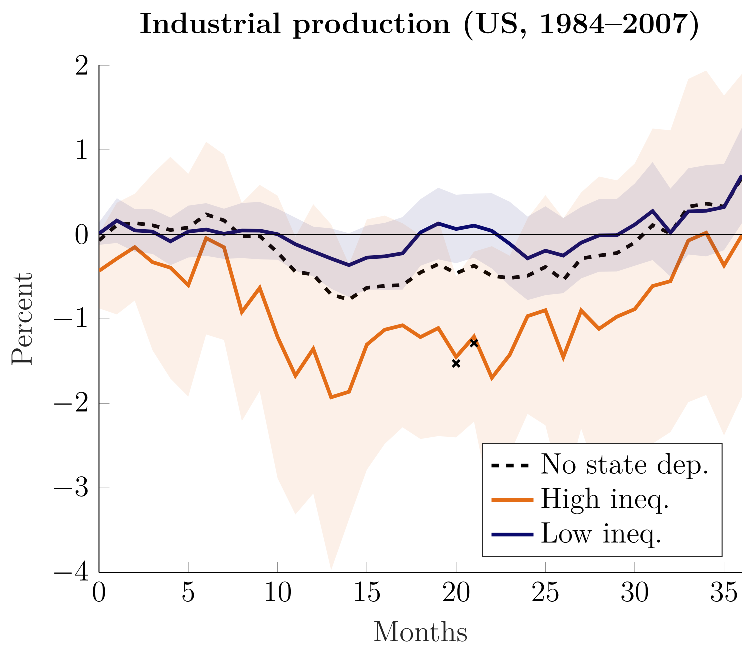


Figure 2.16: Impulse response to a 25 basis points change of the fed. funds rate.
Notes: Sample period: 1984m1–2007m12. Black dashed: No state dependence is imposed. Blue: Regime of low inequality (first quartile of observed inequality). Red: Regime of high inequality (third quartile). Confidence intervals are at the 90% level and based on Newey and West (1987) standard errors. Black crosses indicate horizons at which we reject $H_0 : \beta_h^+ = 0$ at a 90% confidence level.

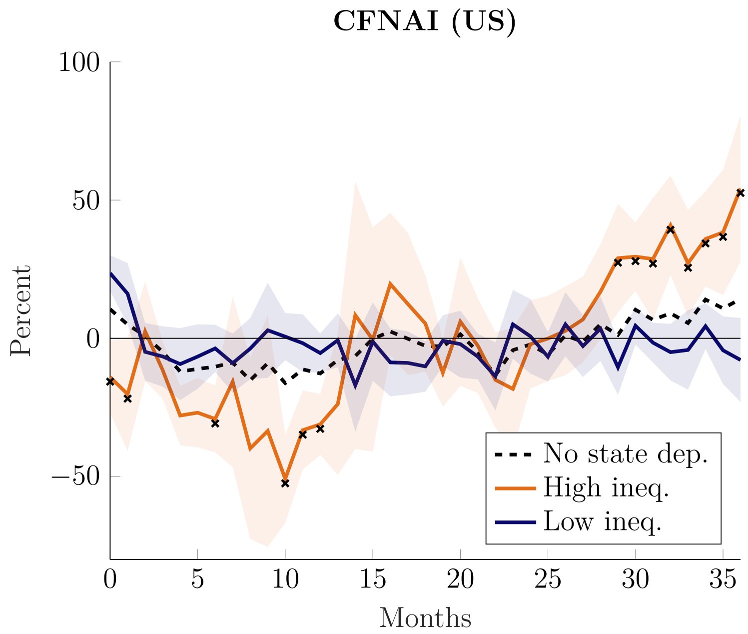


Figure 2.17: Impulse response to a 25 basis points change of the fed. funds rate.
Notes: Measure of output is the Chicago Fed National Activity Index. Black dashed: No state dependence is imposed. Blue: Regime of low inequality (first quartile of observed inequality). Red: Regime of high inequality (third quartile). Confidence intervals are at the 90% level and based on Newey and West (1987) standard errors. Black crosses indicate horizons at which we reject $H_0 : \beta_h^+ = 0$ at a 90% confidence level.

B US States

B.1 Construction of the top 1% wealth share

For each US state, the SOI Bulletins provide estimates of two numbers that we use in our computation, the total wealth held by the top wealth holders and the number of top wealth holders. In some of the years, however, the data provided by the IRS is restricted to a subset of the top wealth holders that lie below or above certain cut-off levels in terms of the gross assets or net worth they possess. Table 2.8 lists these cut-offs.

In a first step, we therefore adjust the given values for total net worth of top wealth holders and the number of top wealth holders to correspond to the given threshold levels in terms of *net worth*. For instance, in 1976 we ask what fraction of the top wealth holders' wealth is owned by the top wealth holders whose net worth exceeds \$120,000. This is only a subset of the top wealth holders whose gross assets lie beyond \$120,000, for which the value is known. The SOI Bulletins provide information on how much net worth is owned by top wealth holders with gross assets exceeding but net worth falling short of \$120,000 on the aggregate US level. We can therefore compute the fraction f of net worth that is owned by these (poorest) top wealth holders. We then multiply the net worth of top wealth holders in each US state by $(1 - f)$ to arrive at the adjusted total net worth of top wealth holders. We proceed analogously for the number of top wealth holders.

For each state and each year for which the IRS provides information on top wealth holders

Table 2.8: Wealth cut-offs in between which estimates of wealth holdings are available.

Year	Cut-off below (≈ 1-percentile)	Cut-off above (≈ 1-percentile)	SOI Bulletin	Author(s) of SOI Bulletin
1976	$120,000 (6.1)	- (0)	Summer 1983	M. Schwartz
1982	$325,000 (2.8)	- (0)	Spring 1988	M. Schwartz
1986	$500,000 (1.6)	$\underline{\$10\text{m}}$ (0.012)	Spring 1990	M. Schwartz & B. Johnson
1989	$600,000 (1.9)	$\underline{\$10\text{m}}$ (0.020)	Spring 1993	B. Johnson & M. Schwartz
1992	$600,000 (2.0)	$\underline{\$10\text{m}}$ (0.020)	Winter 1998	B. Johnson
1995	$600,000 (2.5)	$\underline{\$10\text{m}}$ (0.023)	Winter 2000	B. Johnson
1998	$\underline{\$1\text{m}}$ (1.4)	- (0)	Winter 2003	B. Johnson & L. Schreiber
2001	$\underline{\$1\text{m}}$ (1.7)	- (0)	Winter 2006	B. Johnson & B. Raub
2004	$\underline{\$1.5\text{m}}$ (1.0)	- (0)	Fall 2008	B. Raub
2007	$\underline{\$2\text{m}}$ (0.8)	- (0)	Winter 2012	B. Raub & J. Newcomb

Notes: Numbers are not underlined if they correspond to gross assets and underlined if they correspond to net worth.

we are now equipped with the following information:

- The (adjusted) definition of a top wealth holder: A person is a top wealth holder if her net worth lies between a and b dollars, where a and b correspond to the cut-offs shown in Table 2.8.
- The total net worth of top wealth holders in a state: Denote this number by NW.
- The share of top wealth holders in the population: Denote this number by s.

The goal is to use this information to construct a measure of inequality that is comparable across states. We make the following assumptions.

Assumption 1. *The distribution of net worth x for some state at some time is given by the probability distribution function f. Its right tail is assumed to follow a Pareto distribution, i.e. for some $x_M < a$*

$$f(x) = \begin{cases} \phi(x) & x \leq x_M \\ g(x)\,(1-\Phi(x_M)) & x > x_M \end{cases},$$

where g is the density of the Pareto distribution with scale parameter x_M and shape parameter k and ϕ is some unknown pdf with Φ the corresponding cdf.

Denote by N the total population in a state. It follows

$$\begin{aligned} NW = N \int_a^b x f(x)\,\mathrm{d}x &= N \int_a^b x g(x)\,\mathrm{d}x\,[1-\Phi(x_M)] \\ &= N x_M^k \frac{k}{1-k}\left(b^{1-k} - a^{1-k}\right)[1-\Phi(x_M)] \end{aligned}$$

and

$$s = \int_a^b f(x)\,\mathrm{d}x = (G(b) - G(a))\,[1 - \Phi(x_M)] \tag{2.4}$$

where $G(\cdot)$ is the cdf of the Pareto distribution. The average net worth of a top wealth holder $NW/(N \cdot s)$ implicitly defines the shape parameter k of the Pareto distribution

$$\frac{NW}{N \cdot s} = \frac{\int_a^b xg(x)\,\mathrm{d}x}{G(b) - G(a)},$$

The total wealth of the top T wealth holders, for $T < 1 - \Phi(x_M)$, is given by

$$TopTPercWealth = N \int_t^\infty xf(x)\,\mathrm{d}x,$$

with $t = x_M(\frac{1-\Phi(x_M)}{T})^{1/k}$. This yields

$$TopTPercWealth = N\frac{k}{k-1}x_M T^{\frac{k-1}{k}}\,[1 - \Phi(x_M)]^{\frac{1}{k}},$$

and using (2.4)

$$TopTPerWealth = N\frac{k}{k-1}T^{\frac{k-1}{k}}s^{1/k}\left(a^{-k} - b^{-k}\right)^{-1/k}. \tag{2.5}$$

Dividing (2.5) by the state's total wealth gives the wealth share of the richest $T \cdot 100\%$. As a measure of the total population size N we use data from the U.S. Census Bureau (population aged 20 and above).

Lastly, we require a measure of total wealth on the state level, which is not readily available. We therefore use the following procedure. We obtain total US household wealth that corresponds to the items included in the net worth measures used by the IRS from Kopczuk and Saez (2004).[25] We then use data on capital income from the Bureau of Economic Analysis (item "Dividends, interest, and rent") which is available both for the US and on the state level. We then divide total US net wealth by capital income, backing out an aggregate interest rate. We then use this interest rate to capitalize capital income on the state level, i.e. we multiply capital income on the state level with the aggregate interest rate. Implicitly we therefore assume that the portfolio of assets in each state earns the same interest rate in a given year.

Figure 2.18 plots our self-constructed top 1% share for the US, as well as the estimate from Kopczuk and Saez (2004) (extended by Saez and Zucman (2016) for the years 2001 and 2004) who rely on confidential individual estate tax return data. While our measure displays a somewhat higher level as well as larger amplitude over time than theirs, overall, both the

[25] Their time series on total wealth ends in 2002 and we therefore miss two observations, 2004 and 2007. Since the ratio of total US household net worth as measured by the Federal Reserve Board (series TNWBSHNO) to Kopczuk and Saez (2004)'s measure has historically been very stable (mean: 1.36, max: 1.41, min: 1.32), we divide the Fed's measure by the mean of this ratio to obtain total net worth for the two last years in our sample.

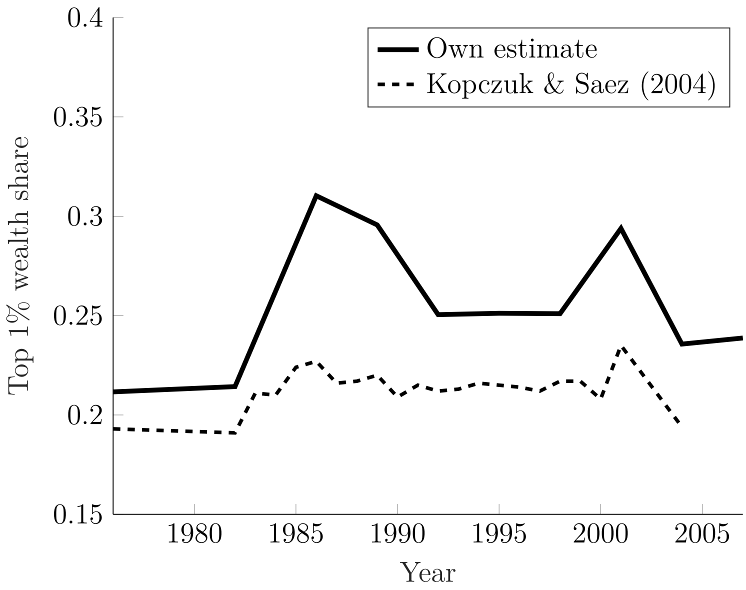


Figure 2.18: Top 1% wealth share in the US.

level as well as the dynamics of our wealth share are broadly consistent with Kopczuk and Saez (2004) and Saez and Zucman (2016).[26] Note also that for the identification of the effects of wealth inequality on the states' responses to monetary policy in Section 3 we only require that we do not make any systematic error in measuring wealth inequality *across* US states. Over- or underestimating the top 1% share at any point in time *in all states* would not confound our results.

Figure 2.19 shows the top 1% wealth share, averaged over time, for all US states. Nevada is the state with the highest estimated average wealth inequality (top 1% share of 35.7%), followed by California, Connecticut, and New York. Both Texas and Florida also feature above-average inequality. In the mid-western states, most notably North Dakota (top 1% share of 14.3%), Iowa, Montana and South Dakota wealth is rather equally distributed. These patterns are broadly in line with the concentration of top wealth holders or millionaires in the population, reported graphically in several of the SOI Bulletins.

B.2 Additional regression tables

[26]The fact that estimates of top wealth shares using the estate tax multiplier method do not show as stark an increase in inequality since the 1980s as do estimates based on capitalizing income (Section 2) or based on the Survey of Consumer Finances is well documented. Kopczuk (2015) and Saez and Zucman (2016) discuss potential reasons for this, among others a rising mortality gradient in age over time and increasing estate tax planning.

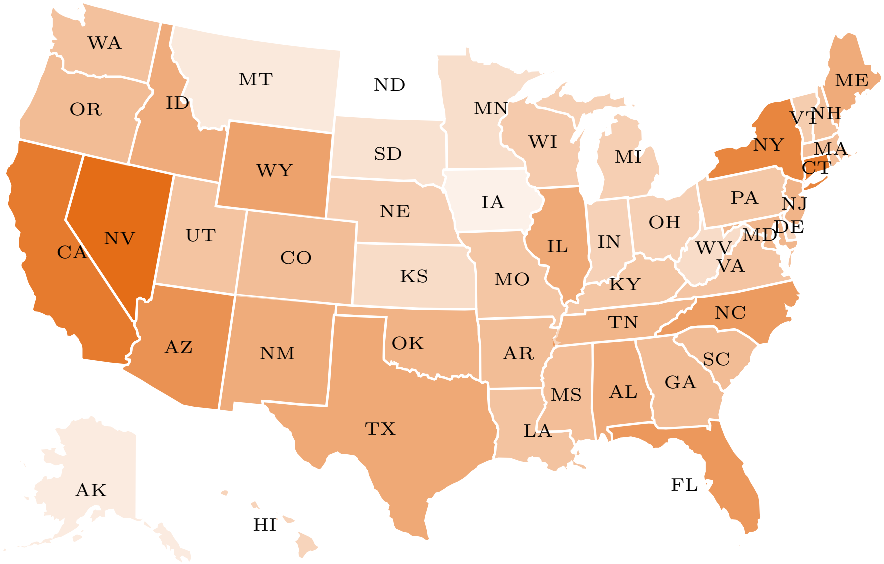


Figure 2.19: Avg. top 1% wealth share 1976–2007, darker = higher.

Table 2.9: Summary statistics for variables on the US state level (1976q1 to 2007q4).

	mean	sd	min	max
Avg. response unemployment (in p.p.)	0.08	0.04	0.01	0.17
Top 1% wealth share	0.24	0.04	0.14	0.36
Manuf. share	0.18	0.07	0.04	0.32
Share small firms	0.50	0.06	0.41	0.67
Share middle-aged	0.49	0.01	0.45	0.53
Share old	0.17	0.03	0.07	0.24
Observations	50			

Table 2.10: Regression results for the cross-section of US states (State Coincident Index).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Top 1% wealth share	-0.4	-0.6*	-0.3	-0.2	-1.8**	-1.9***	-1.7**	-1.5*
	(0.3)	(0.3)	(0.3)	(0.4)	(0.7)	(0.6)	(0.8)	(0.8)
Manuf. share		-0.9***	-0.7**	-0.6*		-1.2*	-0.9	-0.7
		(0.2)	(0.3)	(0.3)		(0.7)	(0.8)	(0.8)
Share small firms			0.5*	0.6*			0.5	1.1
			(0.3)	(0.3)			(0.9)	(0.9)
Share middle-aged				-1.1				-5.7**
				(1.0)				(2.3)
Share old				-0.5				-1.5
				(0.6)				(1.3)
Observations	50	50	50	50	50	50	50	50
Sample	'79–'07	'79–'07	'79–'07	'79–'07	'84–'07	'84–'07	'84–'07	'84–'07

Notes: Dependent variable is the cumulative IRF of the log state coincident index (constructed by the Federal Reserve Bank of Philadelphia) to a 25 basis points increase in the interest rate over a horizon of three years. Robust standard errors are reported in parentheses, a * indicates $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.11: Regression results for the cross-section of US states (peak response).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Top 1% wealth share	0.09	0.05	-0.003	0.2	-1.2*	-1.3*	-1.3*	-1.0
	(0.2)	(0.3)	(0.3)	(0.3)	(0.6)	(0.7)	(0.7)	(0.8)
Manuf. share		-0.5***	-0.5**	-0.3		-0.8	-0.8	-0.5
		(0.2)	(0.2)	(0.2)		(0.5)	(0.6)	(0.6)
Share small firms			-0.10	0.1			0.04	0.7
			(0.3)	(0.2)			(0.7)	(0.7)
Share middle-aged				-2.5***				-5.5**
				(0.8)				(2.5)
Share old				-1.0**				-1.5
				(0.4)				(1.3)
Observations	50	50	50	50	50	50	50	50
Sample	'69–'07	'69–'07	'69–'07	'69–'07	'84–'07	'84–'07	'84–'07	'84–'07

Notes: Dependent variable is the minimum response of log real state personal income to a 25 bp increase in the fed. funds rate rate over a horizon of three years. Robust standard errors are reported in parentheses, a * indicates $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.12: Regression results for the cross-section of US states (Fixed effects regressions)

	(1)	(2)	(3)	(4)	(5)	(6)
Top 1% wealth share	-0.6*	-1.8***	-0.8	-0.5	-0.5	-0.6
	(0.3)	(0.6)	(0.6)	(0.5)	(0.5)	(0.5)
Manuf. share				3.0***	2.7***	2.6***
				(0.4)	(0.4)	(0.6)
Share small firms					1.9*	1.1
					(1.0)	(1.4)
Share middle-aged						-1.8
						(1.3)
Share old						5.2*
						(2.8)
Observations	50	50	100	100	100	100
Sample & FE	'69–'88	'88–'07	'69–'07, FE	'69–'07, FE	'69–'07, FE	'69–'07, FE

Notes: Dependent variable is the cumulative IRF of real state personal income to a 25 basis points increase in the interest rate over a horizon of three years. Full sample is cut in half and then a fixed effects regression is conducted (columns 3–6). The first two columns indicate results from estimating equation (2.2) for each of the two subsamples separately. Robust standard errors are reported in parentheses, a * indicates $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Chapter 3

Monetary Policy and Wealth Inequality—The Role of Entrepreneurs

Joint with Alexander Matusche.

1 Introduction

Entrepreneurs, private business owners with a tight connection to their firm, form a relatively small fraction of the total US population, approximately 7.5%. However, this small group owns about one third of total US household wealth. Furthermore, their businesses employ close to half of all US workers.[1] Therefore, entrepreneurs' investment response to an interest rate change might be crucial for understanding the transmission of monetary policy to aggregate employment, investment, and GDP. Using data from the Survey of Consumer Finances (SCF), we document that the gap between the average net worth held by entrepreneurs and by the rest of the population has widened since the 1980s. Put differently, there has been a shift of wealth towards the already wealthy group of entrepreneurs. In this paper, we investigate the quantitative importance of entrepreneurs and the consequences of the observed shift of wealth towards them for the transmission of monetary policy.

As our main contribution, we develop a Heterogeneous Agent New Keynesian (HANK) model in which a fraction of households are entrepreneurs who can decide to invest in private businesses with risky returns. We find that, first, entrepreneurs are quantitatively important for the transmission of monetary policy to the real economy, despite being only a small group of households. Second, a more unequal distribution of wealth, caused by a shift of wealth from workers to entrepreneurs leads to larger output effects of monetary policy.

On the household side of our model, we explicitly distinguish between workers and entrepreneurs. All households, i.e., workers and entrepreneurs, can save in a liquid and an

[1]See Cagetti and De Nardi (2006) and De Nardi et al. (2007) and Section 2 below.

illiquid asset. However, entrepreneurs have access to an additional investment opportunity, their private firm. This firm operates a decreasing returns to scale technology using labor and capital as inputs. Production of the private firm is associated with idiosyncratic risk. Since households cannot trade private firms and markets are incomplete, entrepreneurs cannot insure against this idiosyncratic risk. We merge private firms of entrepreneurs into an otherwise standard New Keynesian supply side.

Of key importance for the transmission of interest rate changes is the resulting portfolio reallocation of entrepreneurs. When monetary policy lowers the interest rate on the government bond, entrepreneurs optimally rebalance their portfolios. They reshuffle parts of their funds into their firm, away from the now lower-yielding bond. Because of decreasing returns to scale in the production function of private firms, an entrepreneur's net worth is a critical determinant of the interest rate elasticity of private firm investment. On the one hand, the elasticity is low for entrepreneurs with little wealth. Returns from investing into their private business are large compared to an investment in the government bond, i.e. they earn high excess returns on private firm investment. An interest rate cut does not increase this excess return much in relative terms, and hence the portfolio reallocation response is small. On the other hand, the investment elasticity is high for wealthy entrepreneurs. On average, they own large firms and thus face a relatively low marginal product of capital within their firm due to decreasing returns to scale. The small excess returns of large firm owners are affected relatively more by the interest rate cut, so their reallocation response is stronger. This portfolio reallocation is a direct effect of monetary policy, i.e., it occurs whenever the central bank changes the interest rate, and it does not require other aggregate variables to be affected.

The poorest entrepreneurs also react strongly to monetary policy, though not because of the just described portfolio reallocation mechanism. They earn high excess returns from investments into their firm so that their marginal propensity to invest is large. Expansionary monetary policy enables them to increase their investment for two main reasons. First, because poor entrepreneurs have strong incentives to grow their firm, they take out debt to grow their business. Expansionary monetary policy makes debt cheaper and thereby generates a positive income effect for this group of entrepreneurs. Second, monetary policy has indirect effects, i.e. in general equilibrium it affects prices other than the interest rate. In particular, expansionary monetary policy stimulates economic activity and generates higher incomes. Poor entrepreneurs use the additional income by investing heavily into their firm. Taken together, the portfolio reallocation effect, the income effect of an interest rate change and the indirect effects of monetary policy result in an elasticity of private firm investment following an interest rate cut that is u-shaped in net worth, i.e. the poorest and the wealthiest entrepreneurs respond most strongly.

We calibrate our model to the US economy. We target the size of the private business sector in terms of employment, as well as the shares of liquid and illiquid assets held by entrepreneurs and workers respectively. Since the distribution of liquid assets crucially affects the transmission of monetary policy in our model, we additionally target the shares of hand-to-mouth workers and of hand-to-mouth entrepreneurs.

We test key implications of the model using data from the SCF. First, we show that the average return that entrepreneurs receive from private firm investment diminishes with increasing net worth. Our model matches the empirical distribution of returns from private businesses very well, both unconditionally and conditional on net worth, even though these statistics are not targeted in the calibration. Second, we provide direct empirical evidence for the portfolio reallocation mechanism of entrepreneurs. Using identified monetary policy shocks, we document that entrepreneurs increase the share of wealth that they hold in firm capital after a decrease in the federal funds rate. The findings also suggest that this response is heterogeneous across entrepreneurs. In line with our model, both entrepreneurs with low and with high returns react most strongly.

The calibrated model implies an important role for entrepreneurs in the transmission of monetary policy, even though they constitute only a small fraction of all households. To show this, we conduct a counterfactual exercise in which we assume that entrepreneurs are ignorant about changes in prices and aggregate quantities induced by a monetary policy shock. Compared to the baseline, in which entrepreneurs factor in changes in aggregates, the responses of output and aggregate investment are muted by about 50% when entrepreneurs are ignorant. It is especially important that entrepreneurs take into account the change in the interest rate on the government bond, because this stimulates private firm investment. This highlights the importance of the direct effects of monetary policy, specifically the portfolio reallocation effect we discussed above.

To understand how differing levels of wealth inequality affect the transmission of a monetary policy shock, we conduct two experiments. We calibrate both experiments such that the share of wealth owned by the top 10% richest households goes up by one percentage point compared to the initial steady state. In the first experiment we compute the approximate aggregate output response to a change in the interest rate for given equilibrium policy functions. We do this once using the actual steady state distribution and once using a counterfactual distribution that exogenously features higher wealth inequality. This exercise has the advantage that we do not need to take a stance on the underlying driver of the increase in wealth inequality. We change the wealth distribution such that the average entrepreneur becomes richer while the average wealth of workers stays unchanged, in accordance with our empirical observations. Under the high-inequality distribution we obtain a 7 to 10% larger output response to an interest rate change, because a larger share of wealth is now held by rich entrepreneurs who react strongly to monetary policy.

Second, we re-parameterize the model such that it endogenously generates a more unequal steady state wealth distribution. We assume that entrepreneurs are already born with relatively large firms, in contrast to our baseline model in which all households begin their lives with zero wealth. This can be motivated by decreasing estate taxation in the US since the 1980s. We find that the output response to monetary policy is amplified by 3 to 20%

relative to the initial economy, depending on how many entrepreneurs are endowed with a positive bequest.

The remainder of this paper is structured as follows. First, we discuss the related literature. In Section 2 we provide empirical evidence that motivates our focus on private business owners. In Section 3 we describe our model, which we calibate in Section 4. In Section 5 we analyze the transmission of monetary policy in the model. In Section 6 we provide empirical evidence on the distribution of entrepreneurial business returns that is consistent with core predictions of our model, as well as evidence from identified monetary policy shocks. We conduct our main experiment in which we investigate the effects of higher wealth inequality on the transmission of monetary policy in Section 7. Section 8 concludes.

Related Literature The importance of entrepreneurs for the US economy has been documented in a number of studies. In particular, Cagetti and De Nardi (2006) and De Nardi et al. (2007) highlight that the average entrepreneur is rich and that entrepreneurs hold about a third of total US wealth. Asker, Farre-Mensa, et al. (2015) estimate that about half of aggregate investment in the US takes place in private firms. Moreover, two recent empirical studies find that entrepreneurs play an important role for monetary policy transmission. First, Bahaj et al. (2020) document that a significant fraction of the aggregate employment response to expansionary monetary policy shocks in the UK is driven by small and medium-sized enterprises, whose owners' collateral constraints relax due to rising house prices. While the precise channel is absent in our model as we abstract from modeling house prices and collateral constraints, this finding demonstrates that business owners feature strong investment responses to interest rate changes, which is in line with our model. Second, Leahy and Thapar (2019) show that US states with a high fraction of middle-aged households display large responses to expansionary monetary policy shocks. They explain this finding with a high likelihood of being an entrepreneur within this age group and therefore stronger effects, because of strongly increasing entrepreneurial activity in a state. To our knowledge, however, our paper is the first to analyze quantitatively the role of private business owners for the transmission of monetary policy in a structural model.

At the same time, a surging literature studies how household heterogeneity affects the transmission of monetary policy, often contrasting heterogeneous agent models with representative agent versions. Early examples include Bilbiie (2008), who demonstrates in a Two Agent New Keynesian model that if constrained agents' income reacts more than one-for-one to aggregate income, there is more amplification of shocks relative to a representative agent model,[2] and Werning (2015), who proves that when income risk and liquidity are acyclical, the aggregate consumption response to an interest rate change in a heterogeneous agent and a representative agent economy coincide.[3] The key distinction between our paper and this

[2]Bilbiie (2020) generalizes this result to HANK models.

[3]Under the more realistic assumptions that income risk is countercyclical and liquidity is procyclical, Werning (2015) finds that the sensitivity of aggregate consumption to interest rates is higher in the incomplete markets model than in the corresponding representative agent model.

strand of literature is that we move beyond a comparison of the polar cases of heterogeneous versus representative agent models. Instead, we ask how in a world with higher inequality, driven by richer entrepreneurs, monetary policy transmission is different than in a world with low inequality.

Kaplan, Moll, et al. (2018) (KMV in all what follows) argue that the distribution of liquid assets matters for the transmission of monetary policy to consumption. While KMV are mostly interested in aggregate consumption, the focus of our paper lies on the aggregate investment response to monetary policy. In addition, on top of workers and a representative firm, which are also present in KMV, we model entrepreneurial households, i.e. private business owners.[4] We emphasize a portfolio reallocation channel that is crucial for the aggregate investment response to monetary policy and its dependence on the wealth distribution. We therefore highlight an important direct effect of monetary policy that is absent if one abstracts from the portfolio choice of entrepreneurs.

We share the focus on the response of aggregate investment to monetary policy in HANK models with a few recent papers and contribute to a further understanding of the investment response by explicitly modeling private business owners. Luetticke (2021) highlights heterogeneity in marginal propensities to invest (MPI) among households and argues that they are high among wealthy individuals. We focus more on the direct effects of monetary policy, in particular the portfolio reallocation following interest rate changes. Auclert, Rognlie, et al. (2020) demonstrate that while indirect effects of expansionary monetary policy are sizable in general equilibrium, it is the investment decision of firms that sets in motion the feedback loop between higher output and larger consumption of households with high marginal propensities to consume (MPC). Bilbiie et al. (2020) demonstrate that models which include both heterogeneous households as well as capital feature an important amplification mechanism caused by income and capital inequality.

Like us, Melcangi and Sterk (2020) study how the wealth distribution affects the transmission of monetary policy transmission. While they focus on stock market participation and portfolio reallocation towards mutual funds as a transmission mechanism of monetary policy, we emphasize the role of private firm investment. Since stock market participation has gone up in the US over time, Melcangi and Sterk (2020) arrive at the same conclusion as we do, namely that monetary policy shocks today have greater effects than in the 1980s.

Lastly, we relate to recent papers by Cloyne, Ferreira, et al. (2020), Jeenas (2019), and Ottonello and Winberry (2020) who link the aggregate consequences of monetary policy shocks to the investment activity of heterogeneous firms. These authors concentrate on publicly listed firms, while we focus on privately owned businesses, whose investment decision arguably has a much tighter connection to its owner's balance sheet. The experiment we conduct, changing the distribution of household wealth, is similar to the one in Ottonello and Winberry (2020). They investigate the effects of changing the net worth distribution of

[4]We follow KMV in using the methods developed in Achdou et al. (2020) to solve our model, which is framed in continuous time.

firms to one in which default risk is higher because firms hold less net worth. In the spirit of our own results, they find that this dampens the effects of monetary policy shocks. We view our paper as establishing a connection between these studies on heterogeneous firms and monetary policy to the literature cited above that has stressed the importance of household inequality for the transmission mechanism of monetary policy.

2 Entrepreneurs in the US

In this section we demonstrate the importance of entrepreneurs, i.e. private business owners, for the US economy. We rely on data from the SCF, using 13 waves of the survey between 1983 and 2019. The SCF oversamples wealthy households, which is a crucial benefit for our analysis compared to other publicly available data sources.

We follow Cagetti and De Nardi (2006) and De Nardi et al. (2007) in defining a household as an "entrepreneur" if the household meets all of the following three criteria:

1. The household head is self-employed
2. The household head owns, or at least partly owns, a private business
3. The household head has an active management role in the business

When applying this definition, we find that only a small share of the population, about 7.5%, qualify as entrepreneurs. Moreover, this share has been rather stable over time.[5] Up until 1992 the public-use SCF files provide detailed information on the industry of the entrepreneur's firm and we list the share of firms in different industries in 1992 in Appendix A. Typical examples of the entrepreneurial firms in our sample include law firms, medical practices, architect's or accounting offices, and firms in construction services, retail and wholesale business. The small group of entrepreneurial households plays a disproportionate role for several aggregate statistics in the US, as we will document next.

Net worth The average entrepreneur is wealthy. As has already been pointed out by Cagetti and De Nardi (2006), entrepreneurs hold about 33% of total US net wealth. Figure 3.1 documents the share of entrepreneurial households in different parts of the US net worth distribution in 2019. While in the bottom 40% of the net worth distribution only about 1.5% of households qualify as entrepreneurs according to our definition, among the top 10% more than every fourth household does, and among the wealthiest one percent every second household owns and manages a private business.

Another way to express the fact that entrepreneurs are relatively rich is to consider how much wealthier the average entrepreneur is compared to the average non-entrepreneur. We plot this ratio in Figure 3.2. Historically, entrepreneurs have been four to eight times

[5] We provide additional graphs, e.g. on the share of entrepreneurs and the aggregate share of wealth held by entrepreneurs over time, in Appendix A.

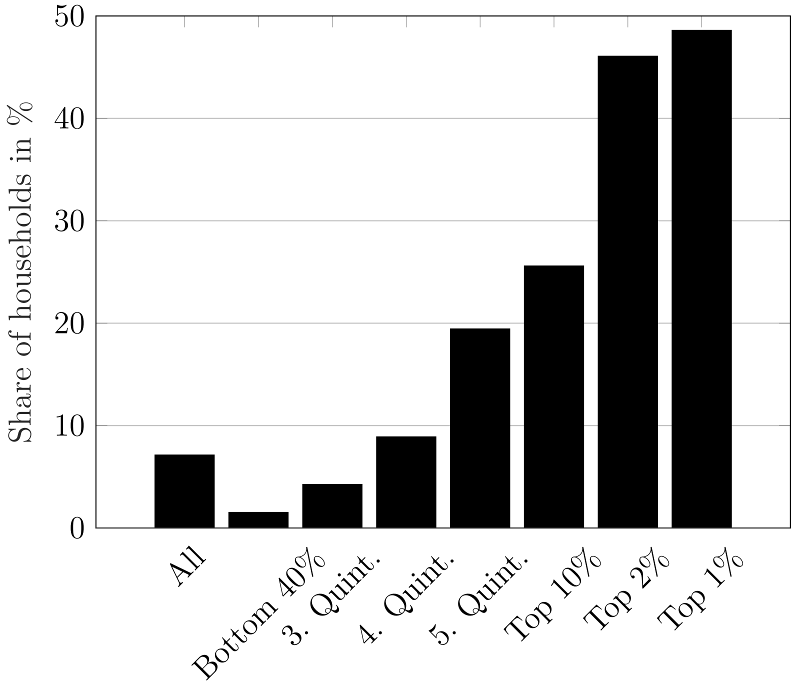


Figure 3.1: Entrepreneurs as a share of all households in different net worth percentiles (SCF 2019).

wealthier than non-entrepreneurs, highlighting again that they are a rich subgroup of the population. Moreover, this ratio has trended upward over time, implying that while entrepreneurs have always been richer than workers on average, the gap between the average entrepreneur and non-entrepreneur in terms of wealth has been widening. To assess whether the time trend is statistically significantly different from zero we conduct a t-test on its coefficient and obtain a p-value of 0.1%. When we exclude the data point in 1983, which could be driving the positive slope, the re-estimated trend coefficient becomes smaller but remains significant with a p-value of 1.0%.

The fact that (rich) entrepreneurs have become even richer compared to the rest of the population is not surprising, given that wealth inequality in the US has been increasing since the 1980s, as documented, for instance, by Kuhn et al. (2020) and Saez and Zucman (2016).[6] In addition, however, wealth has also become more unequally distributed within the group of entrepreneurs. Figure 3.16 in the appendix plots the share of total wealth of entrepreneurs held by the wealthiest 10% of entrepreneurs and shows that inequality has gone up over time. In sum, not only has wealth shifted from non-entrepreneurs to entrepreneurs over the recent decades but it has shifted especially towards the richest ones among them.

Employment To further gauge the significance of entrepreneurs and their private businesses for the US economy, we next demonstrate how important these businesses are for aggregate employment. Starting in 1989, the SCF asks entrepreneurs how many people they

[6]Figure 3.15 in the Appendix documents the trend of rising wealth inequality in the US, as measured by the share of wealth held by the richest 10% of the population.

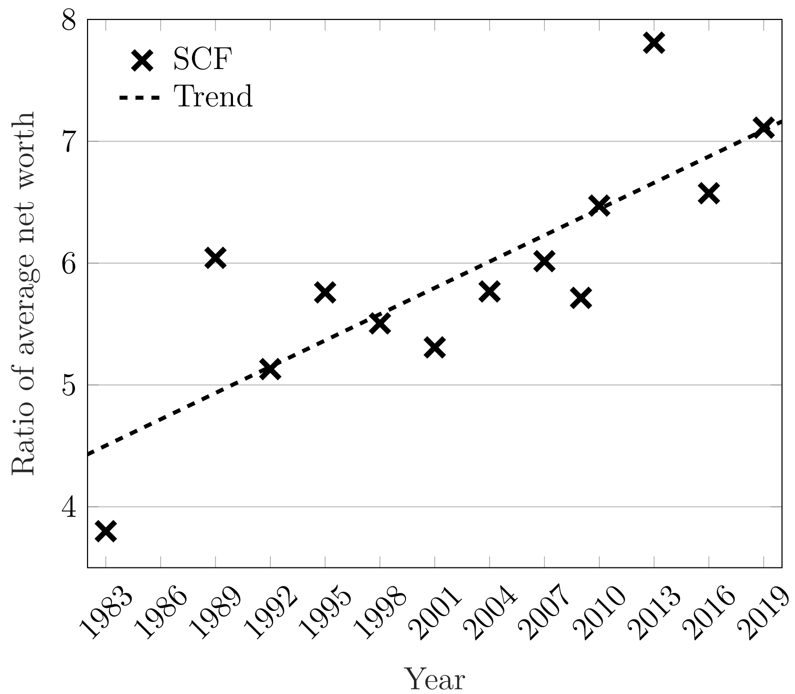


Figure 3.2: Average wealth of entrepreneurs divided by average wealth of non-entrepreneurs.

employ in their businesses. The resulting numbers should be considered as a lower bound of the total employment numbers in the entrepreneurs' firms for two reasons. First, entrepreneurs are only asked about the employment numbers in their businesses for the first two businesses that they own, hence we do not account for employment in any further businesses of the household.[7] Second, the data on employment in privately held businesses is top coded at 5,000 in the public-use SCF files.[8] Another data limitation is that no information on the intensive margin, i.e. hours worked, is given. Instead, households are merely asked how many people they employ in their business. We do not expect our employment figures to overestimate the true numbers because of this issue, as part-time workers predominantly work in industries in which not many of our entrepreneurial businesses are active. We discuss this further in Appendix A.

With these caveats in mind, Figure 3.3 depicts employment in the firms owned by the entrepreneurs in our sample, expressed as a share of total US employment. This share is large, about 46% on average between 1989 and 2019, and similarly to the average wealth ratio displays an upward trend over time (p-value of time trend: 1.5%).[9] While in the late 1980s and early 90s the entrepreneurs' firms contributed to roughly 40% of US employment, this share has risen to approximately 55% in recent years. The time series displays somewhat more volatility than that of the average wealth ratio in Figure 3.2. This is mostly due to the fact that for aggregating to the overall employment share we multiply the average

[7] About 6% of households that we classify as entrepreneurs in the 2019 SCF own more than two firms.

[8] The exact values of the top coding vary over time. While in 1995 to 2019 the upper bound reported in the public files of the SCF is at 5,000 employees, this number is 2,500 in 1989 and 25,000 in 1992.

[9] If we assume that entrepreneurs who own more than two businesses employ as many workers in all their additional businesses as they do in their second business, the employment share is 51% on average.

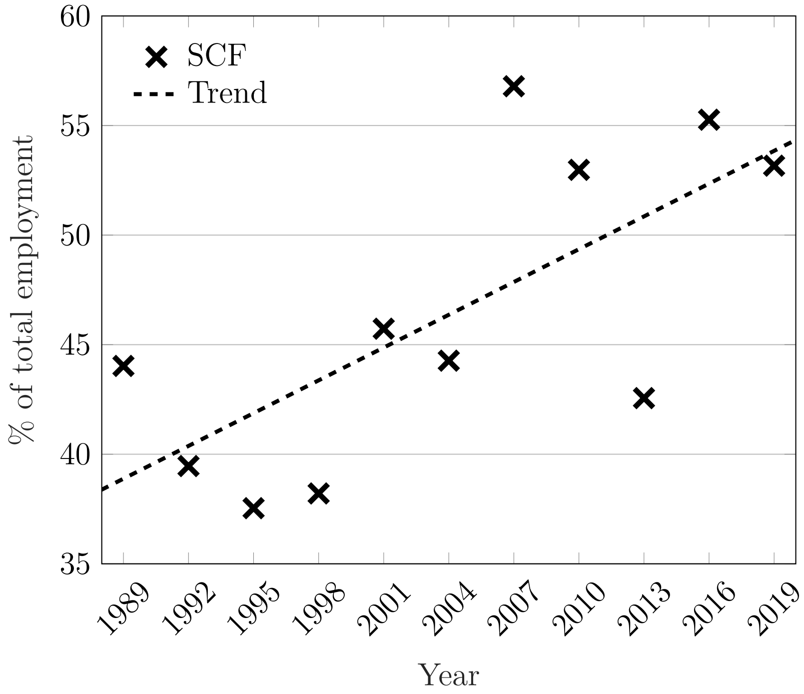


Figure 3.3: Employment in private firms owned by entrepreneurs divided by total US employment.

employee number we obtain from the SCF with the share of entrepreneurial households in the population (Figure 3.12), which itself displays some volatility.[10] Figure 3.14 in the Appendix shows that the average employee number is much less volatile and displays a clear upward trend.

Investment Unfortunately, the SCF contains no reliable information concerning how much the entrepreneurs invest into their private firms. Asker, Farre-Mensa, et al. (2015) estimate that about 53% of aggregate US investment stems from private firms. By focusing on our relatively restrictively defined group of entrepreneurs, however, we are considering only a subset of all private firms in the US, so this figure should be considered as an upper bound of investment undertaken by the firms in our sample. Similarly, when comparing total US aggregate investment to data on capital expenditure from Compustat which captures only the publicly listed firms (as, for instance, in Gutiérrez and Philippon (2017)), about 40% of aggregate investment is left unexplained and would hence be attributable to private firms.

Firm heterogeneity The average numbers on employment mask significant heterogeneity among entrepreneurial firms, whose distribution is heavily skewed. Put differently, there exist many firms that are very small and a small portion of firms that are very large, both in terms of employment and in terms of sales. Appendix A reports more detailed statistics on the firm

[10]In particular, we multiply the average number of employees in the entrepreneurs' firms (corrected for the share of ownership in the respective business) with the share of entrepreneurial households in the population (Figure 3.12), and then multiply this by total households (TTLHH) divided by employment level (CE16OV). All time series used here are obtained from the Federal Reserve Economic Database.

size distribution in our sample. It also shows how the entrepreneurial firms in our sample are distributed across different legal statuses, sources of funding, and industry.

3 A New Keynesian Model with Entrepreneurs

In this section we describe our model, which closely follows KMV's except for the explicit modeling of entrepreneurs. Since we are interested in how the effects of monetary policy shocks change when the degree of wealth inequality in the population varies, the model features heterogeneous households with a realistic distribution of wealth and a New Keynesian supply side (i.e. a HANK model).

Time t is continuous and runs forever. Our model features two types of households, workers and entrepreneurs.[11] Workers are subject to uninsurable labor income risk and work either for the private firms owned by entrepreneurs or for a representative firm that stands for all publicly listed companies. Entrepreneurs have access to a private production technology, and employ their own capital as well as workers, whom they hire on the labor market, to produce output. Input goods are produced both by the private firms and by the representative firm. These input goods are then transformed into intermediate goods by firms that are subject to a price adjustment cost. The intermediate goods firms (and the profits they generate) are owned by the households. The intermediate goods are bundled into the final consumption good, which is then sold to the households. The government consists of a fiscal authority, which levies taxes on households and distributes transfers to them, and a monetary authority which controls the nominal interest rate. All risk in our model is of idiosyncratic nature, i.e. there is no aggregate risk. The monetary policy shock we consider later on is a one-time, unexpected ("MIT") shock.

3.1 Households

At any point in time the economy is populated by a unit mass of households. An exogenous mass s_e of these households are entrepreneurs, and mass $(1-s_e)$ are workers. Each household dies stochastically at rate ζ and is then replaced by a newborn household. These newborn households start their lives with zero assets,[12] a draw from the stationary distribution of the productivity process, and as the same type (worker or entrepreneur) as the household they are replacing.

All households value consumption c_t and dislike labor ℓ_t in the same way. Their preferences are time separable and households discount the future at rate ρ. Taking into account the constant dying intensity ζ, households' preferences over consumption-labor processes

[11] We assume that the occupational choice is exogenous and therefore that household types are fixed over their lifetime, mainly for computational tractability.

[12] This assumption is made for simplicity and follows KMV. In Section 7 we relax this assumption, generating higher wealth inequality in steady state by assuming that entrepreneurs are born with positive assets.

$\{c_t,\, \ell_t\}$ are given by the utility function

$$U\left(\{c_t,\, \ell_t\}\right) = \mathbb{E} \int_{t=0}^{\infty} e^{-(\rho+\zeta)t} u(c_t, \ell_t) \mathrm{d}t, \tag{3.1}$$

where the felicity function $u(c,\ell)$ is additively separable in consumption and labor, monotonically increasing in c and monotonically decreasing in ℓ. It is further strictly concave in both arguments and satisfies $\lim_{\ell \to 0} u_\ell(c,\ell) = 0$ and $\lim_{c \to 0} u_c(c,\ell) = \infty$.

We now turn to a detailed description of the households that are entrepreneurs. Afterwards, we focus on the workers, who face the same problem as in KMV.

Entrepreneurs

Entrepreneurs can invest in three assets. The first is a liquid asset b_t, the second is an illiquid asset a_t, and the third is their private firm of size k_{et}. While the first two are risk-free, investment into the private firm is risky.

We think of the liquid asset as cash and directly held government bonds. The illiquid asset captures houses (net of mortgages), shares in publicly traded firms and pension accounts. Investment in the liquid asset is costless and offers the risk-free return r_t^b. Investment in the illiquid asset is costly. When depositing or withdrawing d_t from the illiquid account of size a_t, the household has to pay a portfolio adjustment cost $\chi^a(d_t, a_t)$. We denote by r_t^a the interest rate earned on the illiquid account. Borrowing is only possible in the liquid asset and only up to a borrowing limit $-\underline{b}$. The interest rate on negative liquid asset holdings exceeds the rate on positive holdings by a constant borrowing wedge κ

$$r_t^b(b_t) = r_t^b + \kappa \cdot \mathbb{1}\{b_t < 0\}.$$

The household cannot hold a negative position of the illiquid asset, $a_t \geq 0$. Due to the adjustment costs, households are only willing to invest in the illiquid account if it yields a higher return, such that in equilibrium we will have $r_t^a > r_t^b$.

In addition to these two assets, entrepreneurs can invest into their own private firm $k_{et} \geq 0$, whose shares are non-tradable. If an entrepreneur wants to grow or shrink her firm she has to pay capital adjustment costs $\chi^e(f_t, k_{et})$, where positive f_t denote enlarging and negative f_t shrinking the firm. Hence, we think of private business capital as a second illiquid asset in the economy. What distinguishes the illiquid asset a from private firm capital k_e is the associated risk. While investment into a is risk-free, investment into the private firm is risky. We specify the sources of this risk after describing the entrepreneurs' production technology.

In order to produce output, entrepreneurs hire labor n_{et} from workers, whom they pay the real wage w_t. The amount of invested capital k_{et} together with the household's productivity and hired labor then determines production of the entrepreneur according to the decreasing returns to scale production function

$$y_e(y, k_e, n_e) = Z_e \cdot y \cdot \left(k_e^{\alpha} \cdot {n_e}^{1-\alpha}\right)^{\nu},$$

with $\nu \in (0,1)$. The parameter $Z_e > 0$ governs the productivity of the entrepreneurial sector relative to the representative firm, whose productivity we normalize to one and whom we describe in more detail further below.

The assumption of decreasing returns is common in the literature on entrepreneurship (Cagetti and De Nardi, 2006; Tan, 2020). It is of key importance for the portfolio reallocation mechanism that we will emphasize later on. The assumption goes back to Lucas (1978) who motivates it using diminishing returns on span-of-control. The entrepreneur's ability in managing the firm gets stretched out over ever larger projects, and accordingly, the productivity of the firm suffers.[13] We provide empirical evidence that wealthier entrepreneurs earn lower returns from their firm in Section 6.1.

We assume that there are two sources of idiosyncratic investment risk. The first source is productivity risk. Current productivity of an entrepreneur y_t evolves stochastically according to some process

$$\dot{y_t} = \Phi_y(y_t).$$

The second source of risk is a capital quality shock that affects the capital employed in the firm k_{et}. We assume that firm capital evolves over time according to the following process:

$$\mathrm{d}k_{et} = [f_t - \delta \cdot k_{et}]\,\mathrm{d}t + \sigma_k \cdot k_{et} \cdot \mathrm{d}W_t \ ,$$

where W_t is a Wiener process, σ_k the standard deviation of the capital quality shock and δ denotes depreciation.

For simplicity, we assume that entrepreneurs themselves work an exogenously fixed amount of hours, $\bar{\ell}$, on tasks regarding the management of the firm, i.e. their work input does not enter the production function.[14] They are not paid wages for this work, compensation for their effort is included in the profits that they receive from their firm. Denoting by p_t the real price of output produced by entrepreneurs at time t, we can define entrepreneurial profits before taxes as

$$\Pi_e(k_{et}, y_t) = p_t \cdot y_e(k_{et}, n^*_{et}, y_t) - w_t \cdot n^*_{et} \ .$$

Here, we have already substituted in the optimal labor demand of the entrepreneurs n^*_{et},

[13] An alternative motivation for arriving at decreasing returns to scale in revenues is to assume constant returns to scale in production and a downward-sloping demand curve for the entrepreneur's output y_e (Asker, Collard-Wexler, et al., 2014; Cooley and Quadrini, 2001).

[14] The precise number of hours worked by the entrepreneurs, $\bar{\ell}$, is irrelevant in all what follows, as utility is additively separable in consumption and labor.

which is a static decision and given by

$$n_{et}^{*} = \left(\frac{p_t \cdot (1-\alpha) \cdot \nu \cdot Z_e \cdot y_t \cdot k_{et}^{\alpha\nu}}{w_t} \right)^{\frac{1}{1-\nu(1-\alpha)}} .$$

Taken together, entrepreneurs maximize utility solving the problem

$$\begin{aligned} &\max_{\{c_t, b_t, d_t, f_t\}} U\left(\{c_t, \bar{\ell}\}\right) \\ \text{subject to:} \quad &\dot{b}_t = (1-\tau_e) \cdot \Pi_e(k_{et}, y_t) + r_t^b(b) \cdot b_t - d_t - f_t + T_t - c_t \\ &\qquad - \chi^a(d_t, a_t) - \chi^e(f_t, k_{et}) + \tau_e \cdot \delta \cdot k_{et} \\ &\dot{a}_t = r_t^a \cdot a_t + d_t \\ &\dot{k}_{et} = f_t - \delta \cdot k_{et} + \sigma_k \cdot k_{et} \cdot \dot{W}_t \\ &b_t \geq -\underline{b},\ a_t \geq 0,\ k_{et} \geq 0, \end{aligned} \tag{3.2}$$

given initial conditions. Here, T_t denotes a lump-sum transfer from the government. The proportional tax on business income τ_e only pertains to profits after depreciation, which gives rise to the tax deduction term $\tau_e \cdot \delta \cdot k_{et}$.

Note that we understand the interest rate on each of the three assets as implicitly augmented by ζ. This accounts for the fact that accidental bequests from the dying households are distributed to the living households in proportion to their current assets (i.e. we assume perfect annuity markets).

Firm dynamics As occupational choice is exogenous, there is no endogenous entry and exit of firms in our model. Hence we also abstract from this margin when we later study the response of the economy to a monetary policy shock. The reader be reminded, however, that households die with probability ζ and are then replaced by households of the same type with zero assets in our model. Hence, exogenous entry and exit exists in our model, and we observe both very large and very small firms in equilibrium.

Workers

As entrepreneurs, workers can invest in the liquid asset b, and the illiquid asset a, but unlike entrepreneurs they cannot run private firms. Instead, they earn labor income and make a continuous labor supply decision on work hours $\ell_t \in [0, 1]$. They supply this labor to the representative firm or to the private firms and are indifferent between these two options, as they receive the same wage (net of a proportional labor tax τ_l) in both cases. Workers receive idiosyncratic shocks to their labor productivity, whose natural logarithm z_t evolves according to some exogenous stochastic process,

$$\dot{z}_t = \Phi_z(z_t).$$

Workers maximize utility solving the problem

$$\begin{aligned} &\max_{\{c_t, \ell_t, b_t, d_t\}} U\left(\{c_t,\, \ell_t\}\right) \\ \text{subject to: } \quad & \dot{b}_t = (1-\tau_l)\cdot w_t \cdot \exp(z_t)\cdot \ell_t + r^b_t(b)\cdot b_t - \chi^a(d_t, a_t) - d_t + T_t - c_t, \\ & \dot{a}_t = r^a_t \cdot a_t + d_t \\ & b_t \geq -\underline{b},\ a_t \geq 0, \end{aligned} \tag{3.3}$$

given initial conditions.

In Appendix C we provide the Hamilton-Jacobi-Bellman equations that characterize the solution to the household problems recursively for the specific process for z_t and y_t described in Section 4 below.

3.2 Production

The economy features a standard New Keynesian supply side with one complication: The first layer of production does not only consist of a representative firm which uses capital and labor to produce an input good. Rather, input goods are produced both by a representative firm and by entrepreneurs (see Figure 3.4). We assume that all entrepreneurs as well as the representative firm produce input goods that are perfectly substitutable. The input goods are then differentiated by monopolistically competitive intermediate goods producers that are subject to price adjustment costs as in many standard New Keynesian models. Lastly, the intermediate goods are sold to a final goods producer, who bundles them and produces the final good which is used for consumption and investment.

The representative firm employs labor N_{pt} and capital K_{pt} to produce output Y_{pt} which it sells at price p_t. The price it gets for selling its output is the same that the entrepreneurs receive because of our assumption of perfect substitutability between output of the representative firm and entrepreneurial production. We assume that the firm operates a Cobb-Douglas production function

$$Y_{pt} = N_{pt}^{1-\alpha} K_{pt}^{\alpha}.$$

Profit maximization requires that factor prices equal marginal products

$$\begin{aligned} r^k_t &= p_t \cdot \alpha \cdot \left(\frac{K_{pt}}{N_{pt}}\right)^{\alpha-1} \\ w_t &= p_t \cdot (1-\alpha)\cdot \left(\frac{K_{pt}}{N_{pt}}\right)^{\alpha} \end{aligned}$$

A continuum of mass one of monopolistically competitive intermediate goods producers buys the general input goods from entrepreneurs and the representative firm at price p_t, and

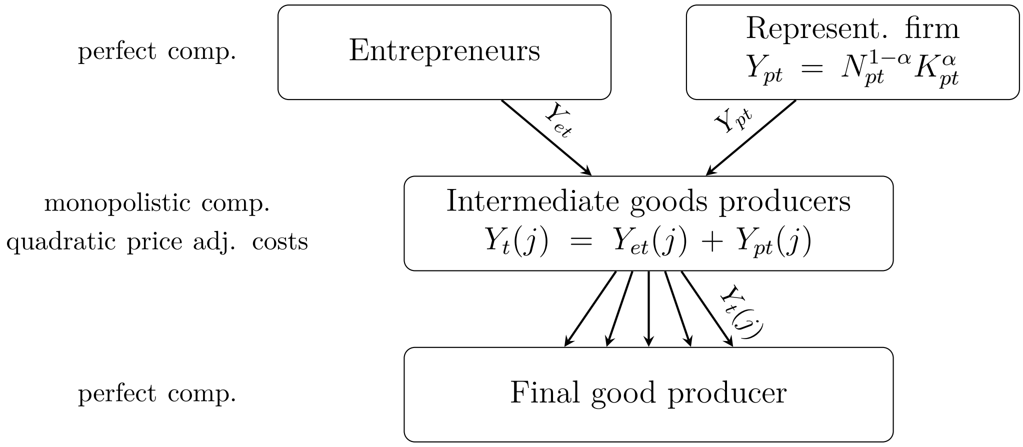


Figure 3.4: Production flow

differentiates them using a linear production technology

$$Y_t(j) = Y_{et}(j) + Y_{pt}(j) \; .$$

Every intermediate good producer j sets the nominal price $P_t(j)$ at which the intermediate good is sold to maximize the present value of real profits. When setting the price, the intermediate good producer takes into account price adjustment costs and the demand schedule $Y^d\left(\frac{P_t(j)}{P_t}\right)$, where P_t denotes the aggregate price level. Price adjustment costs are of the quadratic form as in Rotemberg (1982)

$$\Theta\left(\frac{\dot{P}_t(j)}{P_t(j)}\right) = \frac{\theta}{2} \cdot \left(\frac{\dot{P}_t(j)}{P_t(j)}\right)^2 \cdot Y_t \; ,$$

where Y_t denotes final output and θ is a parameter determining how high the price adjustment costs are.

We assume that the firm discounts the future at rate r_t^a. This is the rate of return of the mutual fund which owns the firm's shares as described below. The maximization problem of an intermediate goods producer is then

$$\max_{\{P_t(j)\}_{t\geq 0}} \int_{t=0}^{\infty} e^{-\int_0^t r_s^a \mathrm{d}s} \left[\left(\frac{P_t(j)}{P_t} - p_t\right) \cdot Y^d\left(\frac{P_t(j)}{P_t}\right) - \Theta\left(\frac{\dot{P}_t(j)}{P_t(j)}\right)\right] \mathrm{d}t.$$

This notation makes it clear that the price of the input goods p_t acts as the real marginal cost of the intermediate goods producers, i.e. $mc_t \equiv p_t$.

The demand schedule is derived from the profit maximization problem of a final good producing firm which combines the intermediate goods into the final output good Y_t according to the production function

$$Y_t = \left(\int_0^1 Y_t(j)^{1-\frac{1}{\epsilon}} \, \mathrm{d}j\right)^{\frac{\epsilon}{\epsilon-1}} \; ,$$

where $\epsilon > 0$ governs the demand elasticity. Profit maximization of the final goods producer yields the demand for intermediate goods

$$Y^d\left(\frac{P_t(j)}{P_t}\right) = \left(\frac{P_t(j)}{P_t}\right)^{-\epsilon} \cdot Y_t.$$

We show in Appendix C that this demand function together with the profit maximization problem of the intermediate goods producer yields the New Keynesian Phillips curve

$$\left(r_t^a - \frac{\dot{Y}_t}{Y_t}\right)\pi_t = \frac{\epsilon}{\theta}\left[mc_t - \frac{\epsilon-1}{\epsilon}\right] + \dot{\pi}_t. \tag{3.4}$$

where $\pi_t = \frac{\dot{P}_t}{P_t}$ denotes the inflation rate. Per period profits net of price adjustment costs are

$$\Pi_t = (1 - mc_t) \cdot Y_t - \frac{\theta}{2} \cdot \pi_t^2 \cdot Y_t$$

3.3 Mutual fund and profits from intermediate goods producers

We assume that households hand their holdings of the illiquid asset a_t to a mutual fund. The fund rents capital K_{pt} at rate r_t^k to the representative firm and invests in shares of the intermediate goods producers, which trade at price q_t. We normalize the total number of shares to one. Optimality of the portfolio allocation requires that the returns on both investments are the same,

$$\frac{\omega \cdot \frac{Y_{pt}}{Y_t} \cdot \Pi_t + \dot{q}_t}{q_t} = r_t^k - \delta = r_t^a. \tag{3.5}$$

We assume that only a fraction $\omega \cdot \frac{Y_{pt}}{Y_t}$ of the profits is paid out as dividends to the mutual fund, where $\omega \in [0, 1]$ is a parameter. The fraction $(1 - \omega)$ is paid out to the workers as a transfer into their liquid account, in proportion to their current labor productivity.[15] The remaining share $\omega \cdot \left(1 - \frac{Y_{pt}}{Y_t}\right) = \omega \cdot \frac{Y_{et}}{Y_t}$ of profits are paid into the liquid account of the entrepreneurs in proportion to the output of their firm. Splitting up the profits Π_t in this fashion ensures that investment into the private and into the representative firm are similarly affected by movements of the profits following a monetary policy shock.

3.4 Government

The government consists of a fiscal and a monetary authority. The fiscal authority collects taxes on labor income (including the part of profits that is paid into the liquid account of workers) and issues real bonds denote by B^S, which assumes a positive value when the government has debt. It pays out transfers to the households and spends an amount G on

[15] This is analogous to the treatment of profits in KMV, where $Y_p = Y$ as they do not model an entrepreneurial sector.

government expenditures. The government's budget is balanced at each instant

$$G + r_t^b B^S + T_t = \tau_l (w_t N_t + (1-\omega)\Pi_t) + \mathrm{Rev}_{et} \ , \tag{3.6}$$

where N_t denotes aggregate labor supply (i.e. labor N_{pt} supplied to the representative firm plus labor supplied to all private firms), and Rev_{et} denotes revenues from taxing entrepreneurial profits, all defined in Appendix C.1.

The monetary authority sets the nominal interest rate $i_t = r_t^b + \pi_t$. We assume that it follows a Taylor rule

$$i_t = \bar{r} + \phi\pi_t + \epsilon_t, \tag{3.7}$$

where ϵ_t denotes a monetary policy shock. It is zero in steady state. Below we consider the effects of an unexpected change of ϵ_t followed by a return back to zero at rate $\eta = 0.5$

$$\epsilon_t = \exp(-\eta t) \cdot \epsilon_0.$$

3.5 Equilibrium

We denote by μ_{wt} the distribution of workers over the state space (b, a, z), and by μ_{et} the distribution of entrepreneurs over (b, a, k_e, y). Both of these distributions integrate to one at every point in time. Let μ_t denote the joint distribution of the two household types. Wherever convenient, we abbreviate by i the set of idiosyncratic state variables, which also includes information about aggregate variables implied by the distribution μ_t. We relegate the definition of the equilibrium to Appendix C.1.

4 Calibration

Wherever possible, our calibration strategy closely follows KMV. We use the same income process for workers and the same values for the externally calibrated parameters and present them in Table 3.2. We only discuss these briefly in the next subsection. Our strategy for the calibration of the remaining parameters, especially those governing the behavior of entrepreneurs, is described in Section 4.2

4.1 Externally calibrated parameters and functional forms

The utility function is

$$u(c, \ell) = \frac{c^{1-\sigma}}{1-\sigma} - \varphi \frac{\ell^{1+\gamma}}{1+\gamma} \ ,$$

where we set both σ and γ equal to 1, and φ to 2.2. These choices ensure a Frisch elasticity of labor supply of one and an average labor supply of approximately 0.5.

Households die at rate $\zeta = \frac{1}{180}$, which implies an average life span of 45 years. The borrowing limit, $\underline{b}$, is set to the average quarterly labor income. The portfolio adjustment

Table 3.1: Parameters of the Income Process

	β_{zj}	λ_{zj}	σ_{zj}
$j = 1$	0.761	0.080	1.74
$j = 2$	0.009	0.007	1.53

cost function for the illiquid asset a is a convex function as in Alves et al. (2020)

$$\chi^a(d,a) = \chi_1^a \cdot \left(\frac{|d|}{a}\right)^{\chi_2^a} \cdot a \ ,$$

where χ_1^a and χ_2^a are parameters.

The stochastic log productivity process of the workers $\Phi_z(z_t)$ consists of two additive parts, a transitory component $z_{1,it}$ and a more persistent component $z_{2,it}$, where i indexes the (worker) household. Therefore, log productivity is

$$z_{it} = z_{1,it} + z_{2,it} \ .$$

Each of the two components follows a jump-drift process, with jumps arriving at rate λ_{zj}. At all times, the process drifts toward its mean of zero at rate β_{zj}. Whenever there is a jump, a new productivity state is drawn from a normal distribution, with $z'_{j,it} \sim N(0, \sigma^2_{zj})$. Hence we have

$$\mathrm{d}z_{j,it} = -\beta_{zj} \cdot z_{j,it} + \mathrm{d}J_{j,it} \ ,$$

where $\mathrm{d}J_{j,it}$ captures the jumps in the process. The parameters for this process are shown in Table 3.1. This is the same income process as in KMV and we refer the reader to their paper for a more detailed discussion. Importantly, the income process ensures that the variance and the kurtosis of the innovations of the modeled income process correspond to those estimated from social security data.

The share of capital in production α assumes a value of 0.33, capital depreciates at 7% annually. The parameter governing the fraction of profits that is automatically reinvested into the illiquid account, ω, is set equal to α. This, as KMV show, ensures that the effect of cyclical profits on investment is sterilized. The demand elasticity faced by the intermediate goods producers, ϵ is set to 10. A value of θ of 100 then ensures that the slope of the Phillips curve is 0.1. The parameter governing the response of the central bank to inflation, ϕ is set to 1.25. We set the government bond supply such that the steady state interest rate on the liquid asset is 2% annually. The lump-sum transfer to the households is set to 6% of GDP and the tax rate on labor income τ_l to 30%.

4.2 Entrepreneurial sector

We set s_e to 7.5%, the average share of entrepreneurs in the US population over the previous decades. We take the degree of decreasing returns in production for private firms ν from Tan (2020), who estimates a value of 0.79.[16] Since most of the entrepreneurial businesses in the SCF are sole proprietorships, partnerships or S corporations (see Table 3.13 in the appendix), which are all subject to pass-through taxation, i.e. business income is not taxed within the company but reported as personal income, we set $\tau_e = \tau_l = 30\%$. Recent evidence in Acemoglu, Manera, et al. (2020) shows that this is a reasonable approximation for the average tax rate on S corporations and C corporations.

Stochastic productivity process $\Phi_y(y_t)$ We assume that productivity y of the entrepreneurs can take on two values. We interpret the low productivity state, y_l, as being a low-talent entrepreneur, i.e. what the literature refers to as a subsistence entrepreneur (Poschke, 2013). The other state, y_h, captures highly talented entrepreneurs (i.e. opportunity entrepreneurs). We normalize $\mathbb{E}[y] = 1$, keeping in mind that the parameter Z_e captures overall productivity of the entrepreneurial sector.

Transitions between the two states happen stochastically, at Poisson rate $\lambda_{y,lh}$ from low to high, and at rate $\lambda_{y,hl}$ from high to low state. Given our interpretation of the two states, we assume that switches between the two types take place only very infrequently, similar to the persistent component z_2 of the labor productivity of workers, whose jumps KMV refer to as "career shocks". Accordingly, we calibrate the transition intensities between the two states to occur on average every 38 years.[17] We further assume that 12.3% of entrepreneurs are of the low (subsistence) type, a number we take from Poschke (2013). This then uniquely pins down the transition intensities, $\lambda_{y,lh} = 0.04$ and $\lambda_{y,hl} = 0.006$. At the end of this section we verify that the business income process faced by entrepreneurs in our model is comparable to its analogue in the data, as estimated in DeBacker et al. (2018).

Capital adjustment costs We assume that the entrepreneurial capital adjustment cost function is of a standard quadratic form

$$\chi^e(f, k_e) = \chi_1^e \cdot \left(\frac{f - \delta \cdot k_e}{k_e}\right)^2 \cdot k_e \ ,$$

where χ_1^e is a parameter. This specification ensures that replacing depreciated capital entails no adjustment cost for the entrepreneur.

[16] Strictly speaking, Tan (2020) estimates this value using detailed data on private start-ups, i.e. by construction relatively young firms. We also experimented with a value of $\nu = 0.88$, which is the value used by Cagetti and De Nardi (2006), but results are largely unchanged.

[17] This is as often as a jump takes place on average between persistent worker productivity states, i.e. we impose $\pi_l \cdot 1/\lambda_{y,lh} + (1 - \pi_l) \cdot 1/\lambda_{y,hl} = 1/\lambda_{z2}$, where π_l denotes the mass of entrepreneurs of the low type in the stationary distribution.

Table 3.2: Externally calibrated parameters.

Parameter	Value	Description	Source or Target
s_e	7.5%	share entrepreneurs	SCF
ζ	$1/(4 \cdot 45)$	death rate	avg. lifetime 45 years
σ	1	rel. risk aversion	KMV
φ	2.2	labor disutil.	avg. labor time 8h/day
γ	1	elast. labor supply	KMV
$\underline{b}$	avg. qrtl. lab. inc.	borrowing limit	KMV
α	0.33	capital share	KMV
ω	0.33	dividend ratio	$= \alpha$
ϕ	1.25	infl. response	KMV
ϵ	10	intermed. dem. elast.	mark-up 11%
θ	100	price adj. cost	slope of Phill. Curve 0.1
δ	0.07/4	depreciation rate	KMV
τ_l	0.3	labor tax rate	KMV
T_t	$0.06 \cdot Y_t$	lump-sum transf.	KMV
τ_e	0.3	entrepr. tax rate	$= \tau_l$
$\bar{r}$	0.02/4	steady state int. rate	KMV
ν	0.79	decr. returns to prod.	Tan (2020)

Notes: Rates are expressed as quarterly values.

4.3 Calibration targets

This leaves us with eight parameters to be calibrated. We target the ratio of liquid assets to GDP (0.26), the ratio of illiquid assets to GDP (2.92), the fraction of poor hand-to-mouth households (i.e. those with few liquid and no illiquid assets, 10%), and the fraction of wealthy hand-to-mouth households (few liquid but positive illiquid assets, 20%)[18] to pin down the discount rate ρ, the borrowing wedge κ, and the portfolio adjustment cost function parameters χ_1^a, χ_2^a. We take these targets from KMV.[19]

The remaining four parameters are specific to the entrepreneurial sector in our model. These are the parameters governing the productivity of the entrepreneurial sector, Z_e, the productivity gap between low and high talent types, y_h/y_l, the standard deviation of the capital quality shock, σ_k, and the capital adjustment cost function parameter, χ_1^e.

We use the average employment share of 46% that we found in the SCF to pin down productivity in the entrepreneurial sector. We also want to get the portfolio composition of entrepreneurs correct, at least on average, as portfolios and portfolio reallocation following a monetary policy shock are key to our analysis. To this end, we target the share of liquid assets (b) in the US economy that are held by entrepreneurs (average of 22% across all SCF waves), the share of illiquid assets (a, i.e. not counting private firms) held by entrepreneurs

[18] These hand-to-mouth shares include households of both occupations, i.e. workers and entrepreneurs.

[19] In fact, KMV calibrate an additional third parameter in the portfolio adjustment cost function (χ_0^a). Alves et al. (2020) abstract from this third parameter and we follow them in doing so, as this reduces the numbers of parameters we need to calibrate from nine to eight.

Table 3.3: Internally calibrated parameters

Parameter	Value	Description	Target
ρ	0.018	discount rate	Liquid assets to GDP
κ	0.015	bor. wedge	Illiquid assets to GDP
χ_1^a	0.84	portf. adj. costs	Share wealthy HtM
χ_2^a	1.45	–	Share poor HtM
Z_e	2.01	avg. entre. talent	Empl. share in entrep. firms
y_h/y_l	1.86	spread entre. talent.	Share illiq. assets held by entrep.
σ_k	0.12	capital qual. shock	Share liq. assets held by entrep.
χ_1^e	0.50	adj. costs firm	Share HtM entrep.

(also 22%), and the share of entrepreneurs that are hand-to-mouth (16%).[20]

While the identification of any single parameter cannot be traced back to one single target, there still exist tight linkages between our targets and the calibrated parameters. Capital quality shocks occur frequently in our model, and hence entrepreneurs use liquid assets to insure against them. Hence, the share of liquid assets they hold informs σ_k. In contrast, talent shocks occur very infrequently, and thus entrepreneurs insure against these shocks using the illiquid asset a. This makes the share of illiquid assets held by entrepreneurs a useful target to inform y_h/y_l. Lastly, when capital adjustment costs are high, entrepreneurs grow their firm relatively slowly at the beginning of their lives, hence relatively few of them are up against the borrowing constraint $\underline{b}$.[21] If adjusting capital is cheap, growing the firm quickly in the beginning, all the while facing binding borrowing constraint, becomes more attractive. Therefore, the share of hand-to-mouth entrepreneurs informs χ_1^e.

Table 3.3 lists the calibrated parameters. The first four, also calibrated in KMV, are very close to the values that they find. In terms of parameters regarding the entrepreneurial sector, we find that more productive entrepreneurs (y_h) are about twice as productive as the low-productive ones (y_l). There exists considerable short-term income and investment risk for entrepreneurs, as a capital quality shock of one standard deviation implies a 12% lower or higher capital stock. Lastly, the capital adjustment cost parameter χ_1^e is in the same range as the linear component of the portfolio adjustment cost function, χ_1^a.[22] Given these calibrated values, Table 3.4 documents that we hit the eight targets relatively well.

4.4 Untargeted moments

In Table 3.5 we compare the joint distribution of occupation and wealth, which was untargeted in our calibration, between our model steady state and data from the SCF (the

[20] We construct hand-to-mouth shares in the SCF following the procedure in Kaplan, Violante, et al. (2014).

[21] As mentioned before, all entrepreneurs start their lives with a firm of size $k_e = 0$.

[22] Be reminded, however, that in order to conserve on parameters to be calibrated we have set the parameter governing the convexity of the capital adjustment cost function $\chi^e(\cdot)$ to 2. This is higher than the calibrated convexity parameter of the portfolio adjustment cost function, χ_2^a.

Table 3.4: Targeted moments

	$\frac{K}{Y}$	$\frac{B}{Y}$	pHtm	wHtm	Lab. at e.	Liq. e.	Illiq. e.	Htm e.
Data	2.92	0.26	0.10	0.20	0.46	0.22	0.22	0.16
Model	2.65	0.27	0.10	0.20	0.41	0.21	0.24	0.16

Notes: $\frac{K}{Y}$ is the capital to output ratio, $\frac{B}{Y}$ liquid assets to output, *pHtm* and *wHtm* are the shares of households who are poor and wealthy hand-to-mouth respectively, *Lab at e.* is the share of labor at private businesses, *Liq e.* and *Illiq e.* are the shares of liquid (b) and illiquid (a) assets held by entrepreneurs respectively, *Htm e.* is the share of entrepreneurs who are hand-to-mouth.

Table 3.5: Share of entrepreneurs by net worth percentiles, in %.

	All	1.+2. Quint.	3. Q.	4. Q.	5. Q.	Top 10%	Top 2%	Top 1%
Model	7.5	0.9	2.2	4.8	28.8	43.9	60.9	65.9
SCF 2019	7.1	1.5	4.2	8.9	19.4	25.6	46.1	48.6

values for the SCF correspond to those shown in Figure 3.1). Conditional on picking a random household from a given part of the wealth distribution, we match the probability that the household is an entrepreneur relatively well, though we overstate the likelihood of entrepreneurs appearing at the very top of the wealth distribution. For this reason, the ratio of average wealth held by entrepreneurs and by workers is higher (approximately eleven) in our model than it is in the data (six on average). In terms of overall wealth inequality in our model economy we perform relatively well, as Table 3.6 reveals.

We used the parameters of the income process for entrepreneurs, in particular the standard deviation of the capital quality shock σ_k and the spread between the productivity of high- and low-productivity entrepreneurs, to hit aggregate targets in our calibration. It is therefore crucial to ask whether the income process in our model resembles that of actual entrepreneurs in the US economy. Most suited for such a comparison are recent results by DeBacker et al. (2018), who use a large confidential panel of US income tax returns to scrutinize the business income risk faced by households. Their definition of business income includes the sum of income generated from sole proprietorships, partnerships, and S corporations, and refers to the net profit or loss from business operations after all expenses, costs, and deductions have been subtracted. They do not require households to be actively managing or owning a business, as we do, but we still find it worthwhile to compare the results on income dynamics they report to those generated from our model.

Table 3.6: Shares of wealth held by different groups of the net worth distribution, in %.

	Bottom 50%	Top 20%	Top 10%	Top 1%
Model	0.2	92.5	81.5	36.1
SCF 2019	0.1	87.4	76.5	37.2

Table 3.7: Annual transition matrix of business income (Model/Data), in %.

	1. Quint.	2. Quint.	3. Quint.	4. Quint.	5. Quint.
1. Quint.	$\underline{40}$ / 63	$\underline{24}$ / 19	$\underline{14}$ / 8	$\underline{11}$ / 5	$\underline{11}$ / 5
2. Quint.	$\underline{19}$ / 18	$\underline{39}$ / 49	$\underline{23}$ / 21	$\underline{13}$ / 8	$\underline{6}$ / 4
3. Quint.	$\underline{13}$ / 7	$\underline{19}$ / 20	$\underline{33}$ / 47	$\underline{22}$ / 22	$\underline{12}$ / 4
4. Quint.	$\underline{12}$ / 5	$\underline{13}$ / 7	$\underline{19}$ / 18	$\underline{34}$ / 53	$\underline{23}$ / 16
5. Quint.	$\underline{15}$ / 5	$\underline{5}$ / 3	$\underline{10}$ / 4	$\underline{21}$ / 14	$\underline{49}$ / 75

Notes: The table reports the probability, in percent, of being in the row quintile of business income in a certain year, and in the column quintile in the following year. The data (right values) are from DeBacker et al. (2018). We use the values from their Table 1. We delete their first row and column (corresponding to zero earnings), reweigh the remaining entries such that rows sum to one again, and then consolidate deciles into quintiles. Values from our model (left) are based on a simulation of 10,000 entrepreneurs over two years. We measure business income as $[\Pi_e(k_{et}, y_t) - f_t] \cdot \mathrm{d}t + \mathrm{d}k_{et}$, i.e. as profits net of costs and depreciation.

In particular, DeBacker et al. (2018) report how likely it is for households within a given decile of the business income distribution to end up in another decile in the following year. Table 3.7 shows these numbers, as well as the analogous statistics simulated from our model. For better readability we reduce the dimensionality of their matrix from ten deciles to five quintiles. Also, unlike in their dataset, we do not have the status "no business income" for our entrepreneurs, so we reweight their original transition matrix, deleting the zero income state. In sum, we find that the income process in our model is similar to that in the data, though ours features somewhat more volatility. Even in the data, as highlighted by DeBacker et al. (2018), households face substantial fluctuations in business income, represented by relatively small probabilities of staying in the same earnings quintile year-on-year. The immobility ratio, i.e. the average of the diagonal elements of the transition matrix, is 39% in our model and 57% in the data.

The last untargeted moments that we compare to the data are the returns that entrepreneurs receive from investing into their firm. This is important as a core prediction of our model is that poor entrepreneurs earn higher returns on average than wealthy entrepreneurs. We defer an in-depth discussion of this issue to Section 6.1, but we already point out here that this prediction of the model is indeed borne out by the data. We also demonstrate that the model matches the level of returns both unconditionally, and conditional on net worth. We view this as a success, since nothing in our calibration has directly targeted these statistics.

5 Quantitative Analysis of Monetary Policy

We now analyze the response of the economy to an interest rate change. Our focus is on aggregate investment, in particular the investment response of entrepreneurs and how it

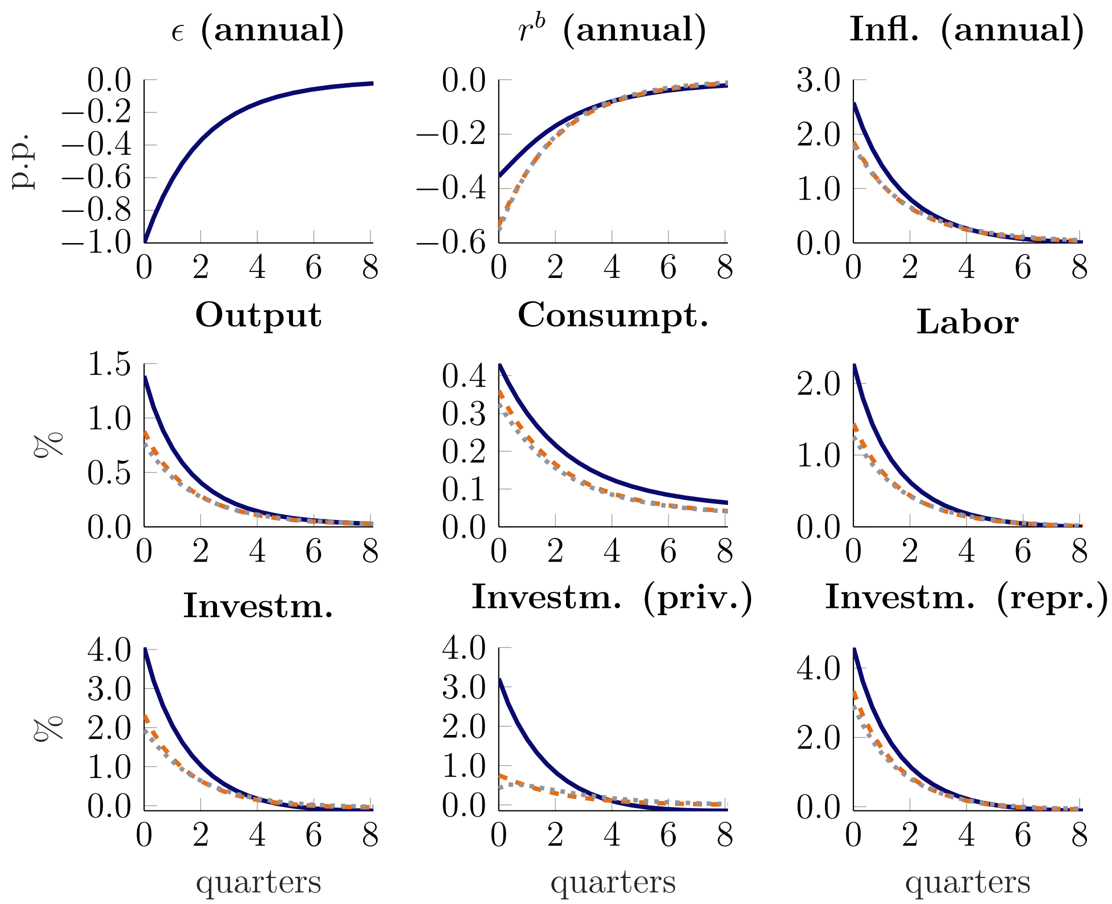


Figure 3.5: Response of aggregate variables to an expansionary monetary policy shock.
Notes: Depicted are deviations from steady state values. The solid blue line shows the model's response to the shock. The dashed orange line shows the response when entrepreneurs (but not workers) are ignorant about changes in all aggregate variables. The dotted grey line shows the response when entrepreneurs are ignorant only about the change in the liquid rate r^b.

depends on the wealth of entrepreneurs and its distribution.

The solid black lines in Figure 3.5 plot the response of the key aggregate variables to an expansionary monetary policy shock. Specifically, we consider an unexpected one time innovation of −100 basis points annually to the Taylor rule (3.7). Owing to the endogenous reaction of the central bank in our model, this leads to a drop in the liquid rate r^b on impact of about 36 basis points. The economy responds in a way that is familiar from the literature on the effects of monetary policy.[23] Output, investment, consumption, labor, and inflation all rise in response to a cut in the interest rate. Output increases by about 1.4% on impact. This number is in line with (though at the upper end of) empirical estimates of the effects of monetary policy shocks (Christiano et al., 2005; Ramey, 2016).[24] Total investment, including that into private firms as well as into the representative firm, rises by 4.1% in our model.

Entrepreneurs, even though they only comprise a small fraction of the total population,

[23]The lack of persistence in the responses to monetary policy shock is a common feature of HANK models. See Auclert, Rognlie, et al. (2020) for a discussion.

[24]Scaling the impulse responses to an impact drop in the real liquid rate r^b of 100 basis points would imply an increase in output of 3.9%.

are important for setting in motion the general equilibrium feedback loop between higher income and higher consumption demand that naturally arises in any HANK model. To see this, we first break down the aggregate output response on impact after the shock into its two most important components, aggregate consumption and investment. Table 3.8 shows that aggregate investment accounts for 56.5% of the overall increase in output while consumption makes up 26.5% (the remaining 17% are accounted for mostly by an increase in price adjustment costs). While entrepreneurs do not contribute significantly to the increase in consumption demand directly, they are responsible for about half of the total increase in investment. Two thirds of their additional investment is directed towards their private businesses, but entrepreneurs also account for about a quarter of the additional investment into the mutual fund and thus capital employed by the representative firm. The expansion in entrepreneurial investment in turn leads to more labor demand and hence higher wages for workers, causing consumption demand to rise further.

Table 3.8: Absolute change in consumption and investment relative to absolute output change (on impact), in %.

	Cons.	Inv. (total)	Inv. (priv.)	Inv. (rep.)
Total	26.5	56.5	17.5	38.9
Workers	24.6	29.5	-	29.5
Entrep.	1.9	26.9	17.5	9.4

Next, we conduct two counterfactual exercises. In both, we assume that entrepreneurs are ignorant about the evolution of a subset of aggregate variables. First, we solve for the dynamics after the monetary policy shock when entrepreneurs do not take into account changes in any of the aggregate variables induced by the shock. We assume that they believe that aggregates are at their steady state values at all times and that they make their consumption and investment decisions accordingly. Importantly, entrepreneurs' assets still evolve according to actual prices. This implies that entrepreneurs still face income changes due to the monetary policy shock and all factor markets clear at all times. In the second case, we assume that entrepreneurs take into account changes in all aggregate variables except of the liquid interest rate r^b, about which they believe that it is constant at its steady state value at all times.[25] This addresses how important the direct effect of monetary policy on entrepreneurs is for the economy's response.

The orange lines in Figure 3.5 correspond to the first experiment in which entrepreneurs are ignorant about all aggregate variables. In this case, the aggregate dynamics following the monetary policy shock are significantly muted. The responses of output and investment to the interest rate change are only about half as large as in the baseline scenario. The investment response is dampened particularly strongly, which highlights again the impor-

[25] The aggregate variables that are relevant for the decisions of entrepreneurs are r^b, r^a, w, p, T, Y_e and Π. The first experiment holds entrepreneurs' believes about all of these constant at steady state values, the second experiment only holds believes about r^b constant.

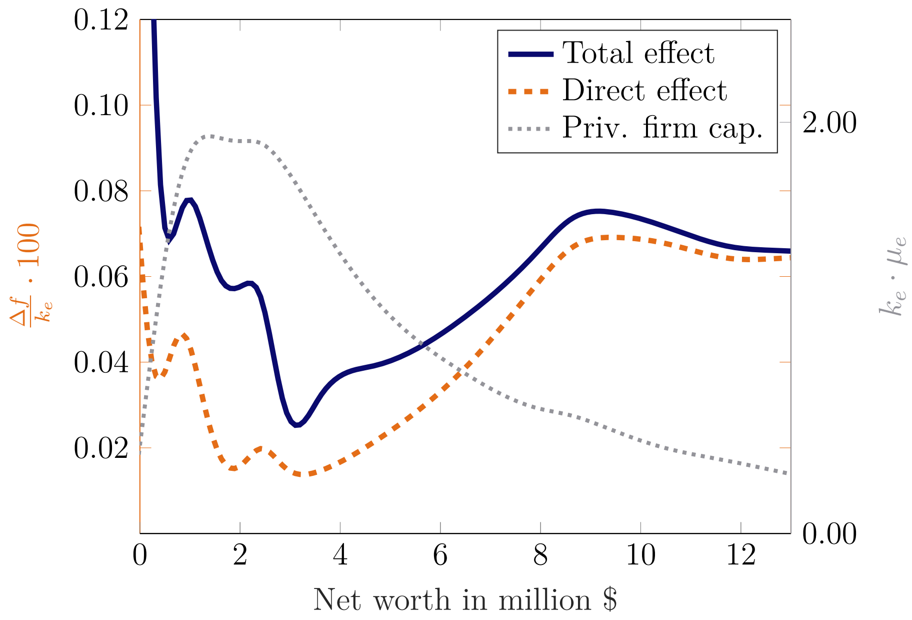


Figure 3.6: Heterogeneous private firm investment response after monetary policy shock.
Notes: Blue solid line: Total change in private firm investment relative to private firm capital in response to monetary policy shock on impact (left y-axis). Orange dashed line: Change in firm investment caused by direct effect, i.e. only liquid rate changes, other prices held fix (left y-axis). Grey dotted line: Share of private firm capital $k_e \cdot \mu_e$ (right y-axis). All lines show averages within small bins of net worth.

tance of entrepreneurs for aggregate investment. Note also that if we were to rescale the impulse responses to imply the same drop in the liquid rate r^b, these results would become even more pronounced. The dotted grey line shows the results for the second experiment, in which entrepreneurs take into account changes in all aggregate variables except that of the liquid rate r^b. The grey lines lie almost on top of the orange lines. Hence it appears crucial to analyze the direct response of entrepreneurs to the change in the liquid interest rate to better understand the transmission of monetary policy in our model. For that, we next turn to the investment responses of entrepreneurs over the net worth distribution.

Heterogeneity among entrepreneurs Entrepreneurs as a group are important for the evolution of aggregate investment in response to the shock, as we have just seen, but there is considerable heterogeneity within the group of entrepreneurs. Depending on the size of their firm and their net worth entrepreneurs respond to the interest rate change very differently.

The solid blue line in Figure 3.6 depicts the total change in the entrepreneurs' private firm investment relative to their capital stock across the net worth distribution. We show the response on impact, i.e. immediately after the expansionary monetary policy shock hits the economy. The line exhibits an approximate u-shape. Relatively poor entrepreneurs respond strongly to monetary policy by expanding their investment. The response is smallest for entrepreneurs with a net worth of around $3 million. For entrepreneurs with net worth above $3 million, the response increases with net worth and then plateaus, so that wealthy entrepreneurs respond relatively strongly.

To better understand what causes this heterogeneity, and guided by the decomposition of the aggregate responses we performed in Figure 3.5, we follow the previous literature on monetary policy shocks in HANK models and distinguish between direct and indirect effects of the interest rate change. Turning to the direct effects first, the orange dashed line in Figure 3.6 depicts the response of firm investment that is due solely to the interest rate change, i.e. keeping all other prices constant. We obtain it by asking how entrepreneurs respond if only the liquid interest rate r^b evolves as it does in Figure 3.5, while all other aggregate variables stay at their steady state levels.[26] The response is smaller than the general equilibrium effect but the overall shape looks very similar. All firm owners invest more into their firm when the liquid rate r^b decreases, optimally reweighting their portfolio and reducing their exposure to the now lower-yielding liquid bond.

The magnitude of this reallocation effect varies greatly with the net worth of a firm's entrepreneur. For the poorest firm owners, at the very left border of the graph, investment elasticities are very large. These are entrepreneurs, who take on debt to grow their firm, as the marginal return from their firm is very large. Once r^b is reduced, they experience a positive income effect as they have to pay lower interest rates on their debt, and hence they invest additional resources into their firm. In terms of aggregate investment, however, these entrepreneurs are of minor importance, as they hold a negligible share of the overall capital stock. This is illustrated by the grey dotted line, which plots the share of business capital held at respective points of the net worth distribution, i.e. $k_e \cdot \mu_e$.[27] A large number indicates that entrepreneurs at this level of net worth hold a large share of total private business capital.

Entrepreneurs with an intermediate amount of net worth do not rely on debt to finance the firm, but their firm still offers relatively high marginal returns. Put differently, the excess return of their firm investment over the riskless bond is large. When monetary policy changes r^b, this excess return is therefore not affected much in relative terms, and hence they rebalance their portfolio relatively little. Wealthy entrepreneurs who own large firms do not reap such high returns from their business. For them, the excess return over r^b is close to zero. Therefore, a change in r^b affects their excess return more significantly, and they reshuffle their portfolio much more than owners of smaller firms to implement the optimal combination of business risk and excess return over the risk-free rate.

Next, we turn to the indirect effects of the interest rate cut on entrepreneurial investment that work through changes in prices and income. The indirect effect on firm investment is the difference between the dashed and the solid line in Figure 3.6, i.e. it is the additional investment that is not due to the interest rate change itself. For most entrepreneurs, the indirect effects are smaller in magnitude than the direct effects. In particular, this is the case for wealthy entrepreneurs. To them the private firm is similar to any other asset in their

[26] In contrast to the decompositions we performed in Figure 3.5, this is a partial equilibrium exercise. We simply ask how entrepreneurs would react to the interest rate path shown in Figure 3.5. We do not ask how prices would have to respond to support the resulting choices as equilibrium outcomes.

[27] The line integrates to the aggregate private firm capital stock $K_e = \int k_e \mu_e(i) di$. To obtain it, we average the private firm capital share within small bins of net worth.

portfolio. Therefore, they spend additional income almost proportionally on investment into the different assets and on consumption, and hence the indirect effect on private business investment is small. For entrepreneurs with little wealth, who own small firms, however, the indirect effects are large. These are households with highly profitable businesses who lack the resources to expand their firm. The rise in income induced by monetary policy allows them to grow their business and they seize this opportunity.[28]

We summarize our results as follows. First, the aggregate output response is affected significantly by the investment decision of entrepreneurs. Second, investment of wealthy entrepreneurs responds more strongly than that of entrepreneurs in the middle of the wealth distribution due to a stronger portfolio reallocation effect. The decreasing returns to scale imply that wealthy entrepreneurs earn a low excess return over the risk-free rate. When the risk-free interest rate falls, they expand investment strongly to restore the optimal excess return. It is therefore crucial that our model produces a realistic distribution of excess returns over net worth. In the next section we look at data from the SCF to argue that it does. Third, poor entrepreneurs also respond strongly to monetary policy because of large indirect effects. However, by definition, they only hold a small fraction of the total capital stock, which mutes their importance for aggregate investment.

6 Implications of the Model: Empirical Evidence

In this section we test two key implications of our model. We do so by relying on data from the SCF, as we did in Section 2.[29] We include households in the sample whose head is aged 25–65, and who have positive net worth and business wealth. Also, as will become clear momentarily, valuations of businesses that are positive but very close to zero result in estimated returns that are very large. We therefore purge the sample of households with the largest 5% of business returns, as defined momentarily.[30] We exclude the SCF wave of 1983 as the variables needed for computing the business returns are only available starting in 1989. Table 3.17 in the appendix reports summary statistics for business returns, net worth and business wealth for the remaining sample over all SCF waves.

6.1 Business returns over net worth

One implication of our quantitative model is that richer entrepreneurs receive smaller returns from their private businesses. To assess this in the data we first need a measure of the business return. Following the baseline definition in De Nardi et al. (2007), we define it as the inverse

[28]Corroborating these results, Figure 3.19 in the appendix plots the marginal propensity to invest into the private business out of a transfer of $500 into the liquid account as a function of net worth.

[29]Wherever necessary, we deflate nominal values to 2019 US dollars using the CPI-U-RS (all items) series obtained from the Bureau of Labor Statistics.

[30]We consider different subsamples below, in particular, we look at the 2019 wave of the SCF and at all waves together. We drop the largest 5% over all waves when we consider all waves and the largest 5% in the 2019 wave when we consider the 2019 wave.

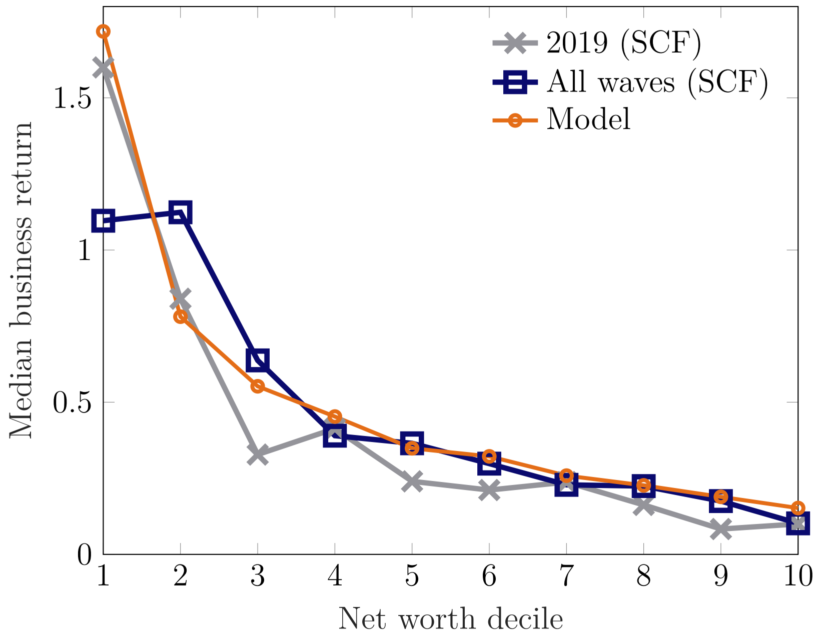


Figure 3.7: Median business returns by decile of the net worth distribution of entrepreneurs.
Notes: Depicted are returns from SCF data only for 2019 (grey), over all waves (blue), and steady state returns from our model (orange).

price to earnings ratio

$$r^e_{it} = \frac{\text{business income}_{it}}{\text{business value}_{it}} .$$

Here, households are indexed by i, the year by t. Business income is the wage or salary income from the main job of the household's head plus business profits paid out to the household (all before taxes).[31] If the head's spouse works at the business we also add wage and salary income of the spouse. For the business value we rely on the answer of households to the question "What is the net worth of (your share of) this business?", i.e. we use the market value of the business.

Figure 3.7 shows the median business return in each decile of the net worth distribution of entrepreneurs for all SCF waves as well as for the most recent wave of 2019. The picture looks very similar for both samples. Returns are substantially lower in higher deciles of the net worth distribution than in lower deciles. Entrepreneurs in the first decile earn an annual return of 110% (all waves) and those in the highest decile of 10%.

The returns for entrepreneurs in the lower net worth deciles appear quite large. One explanation is that the value of their businesses is small and most of the return is actually labor income instead of capital income from their investment. Note that we include wages of the entrepreneur in our definition of business income to be consistent with our model, in which we also do not distinguish between the part of entrepreneurial business income that comes from entrepreneurs' capital investment and the part that comes from their labor input. It is a defining characteristic of entrepreneurship that the two are difficult to distinguish. Lastly, be reminded that all returns shown here correspond to business income before taxes,

[31]By definition of entrepreneurs the head's main job is at the privately owned and managed business.

and hence returns after taxes would be smaller.

We also plot the implied average returns from our model steady state in Figure 3.7, and find that they fit the data surprisingly well, both in terms of levels and in terms of evolution over net worth (the model analogue of net worth is $k_e + a + b$). We view this as a great success of our model to capture a relevant dimension of firm heterogeneity, as nothing in our calibration directly targets either the overall level of returns or the returns conditional on net worth deciles.

The negative relationship between wealth and returns to entrepreneurship is broadly in line with recent empirical evidence in Xavier (2021). Using SCF data as we do, she finds that within the asset class of private businesses returns decline at the top of the wealth distribution. While we observe a largely monotone relationship in the data she discovers an inverse u-shape, with returns largest within the 90th to 97th percentile of the population-wide net worth distribution (in the 2019 SCF, the 90th percentile of the overall net worth distribution corresponds to the 64th percentile of the entrepreneurial net worth distribution, which we use when plotting Figure 3.7). The difference between her results and ours stem from the fact that she does not include the entrepreneurs' labor income in her measure of business profits, which, as we have just argued, tends to raise our measure of profits especially for smaller firms.

The results so far indicate an unconditional correlation between net worth and business returns. In Appendix B we study the relationship between these two variables in more detail. First, we estimate their relationship non-parametrically. Second, we estimate linear regressions in which we control for a large number of observable household characteristics. In both cases a robust negative relationship between net worth and returns emerges. We also discuss potential problems of our analysis and investigate the relationship between business wealth and returns as well as that of business wealth and portfolio shares.

6.2 Portfolio response to monetary policy shocks

To answer how entrepreneurs adjust their portfolios in response to monetary policy, we would ideally observe a panel of entrepreneurial households, preferably at quarterly or even higher frequency, and trace their reaction to identified monetary policy shocks. Unfortunately, the SCF is neither a panel nor does it feature such high frequency, as the data is only collected every three years. We therefore follow an approach similar to Luetticke (2021) who faces the same challenges as we do.

First, for each wave of the SCF we estimate the portfolio share of firm capital, i.e. the ratio of business value to net worth, for each percentile p of the business return distribution and denote the log of this portfolio share by $\gamma_{p,t}$, where t denotes the year of the SCF wave. We non-parametrically estimate these portfolio shares using local linear regressions, effectively using information about the portfolio shares in percentile p and in those percentiles that lie close to p to estimate $\gamma_{p,t}$. Appendix D lays out the details of this procedure.

We then use local projections in the spirit of Jordà (2005) to estimate the effect of monetary policy shocks on the estimated portfolio shares. Specifically, to estimate the effects of an interest rate movement at time t on portfolio shares at $t+h$, we use the regression

$$\gamma_{p,t+h} = \alpha + \beta_{p,h} \cdot FF_t + \delta^{Y}_{p,h} \cdot \ln(Y_{t-1}) + \delta^{FF}_{p,h} \cdot FF_{t-1} + u_{t+h} \ , \tag{3.8}$$

where α is a constant, FF denotes the Federal funds rate, Y is GDP, and u is an error term with $\mathbb{E}[u_t] = 0$. The estimate of interest is $\beta_{p,h}$ and it captures the response at horizon h of the log portfolio share in firm capital to a 100 basis point increase in the interest rate at time t for the p'th percentile of the return distribution.

Since the federal funds rate is endogenous, we instrument it using an identified monetary policy shock series, i.e. we estimate IV local projections. Specifically, we use the narratively identified shock series, denoted ϵ^y_t, from C. Romer and D. Romer (2004) which was extended until 2007 by Ramey (2016). As the shock series ends in 2007, we only use SCF waves 1989 to 2007 in this section. In the appendix we document our results when using the shock series from Gertler and Karadi (2015), who exploit high-frequency financial markets data to construct their shocks, which are available from 1990 to 2012.

To convert the monthly shock series provided by Ramey (2016) into an annual series, we follow the approach proposed by Meier and Reinelt (2020) and Ottonello and Winberry (2020). In particular, we attribute a monthly shock fully to our yearly shock only if it takes place in January. If it takes place later in the year, we partly attribute the shock to the current year and partly to the next year. More specifically, we use the monthly series of shocks ϵ^m_t from Ramey (2016) to construct annual shocks ϵ^y_t according to

$$\epsilon^y_t = \sum_{\tau \in \mathcal{M}(t)} \phi(\tau, t) \cdot \epsilon^m_\tau + \sum_{\tau \in \mathcal{M}(t-1)} (1 - \phi(\tau, t-1)) \cdot \epsilon^m_\tau \ ,$$

where $\mathcal{M}(t)$ is the set of months in year t and

$$\phi(\tau, t) = \frac{\text{remaining number of months in year } t \text{ after announcement in month } \tau}{12} \ .$$

Putting more weight on shocks early in the current year and late in the previous year allows us to more reliably inspect the response of portfolio shares "on impact", i.e. for horizon $h = 0$. However, as some of the respondents answered the survey in the early months of a given year and therefore potentially before some of the monthly shocks of that year materialized, the estimates of $\beta_{p,0}$ have to be interpreted with some caution even when using this particular weighting of monthly shocks.[32]

The left panel of Figure 3.8 depicts our baseline estimates of $\hat{\beta}_{p,h}$ for $h = 0, 1, 2$. Consider the orange solid line first. It depicts the estimated portfolio response on impact to an exoge-

[32] We also performed our analysis using simply the sum of all shocks occurring in a given year. The results are very similar.

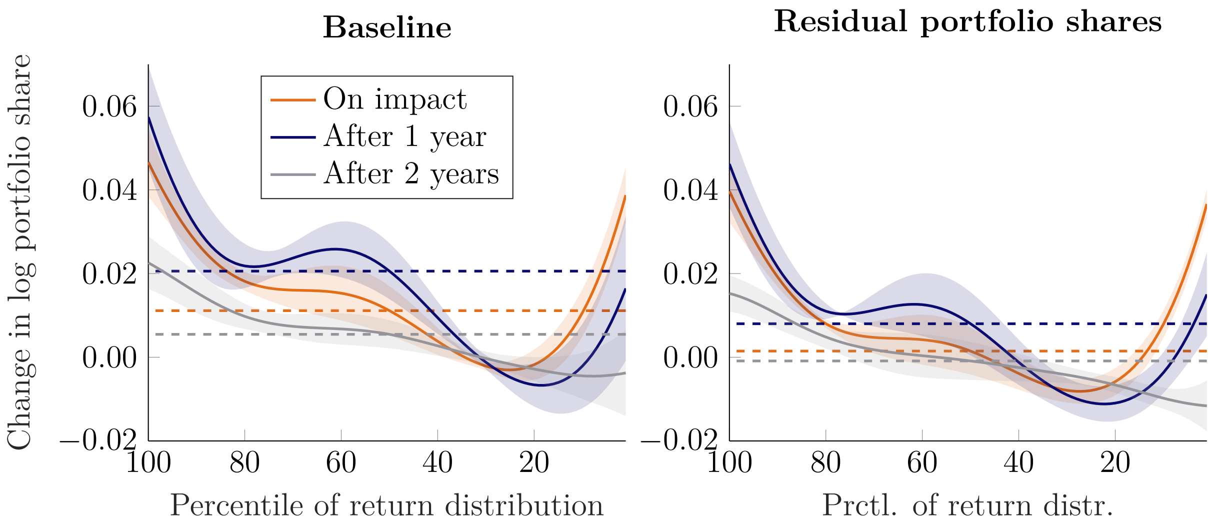


Figure 3.8: Impulse responses of portfolio shares to monetary policy shock.
Notes: Change in the logarithm of portfolio shares following a 25 basis points expansionary monetary policy shock by business return percentile. The dashed lines depict the responses at the median of the return distribution. Confidence bands are at the 66% level.

nous 25 basis point cut in the interest rate for each percentile p of the return distribution, i.e. $-\hat{\beta}_{p,0}/4$. We order percentiles in decreasing order on the x-axis, as we have shown above that high returns typically correspond to the poorest entrepreneurs.

At most percentiles the response is positive, lending evidence to the portfolio reallocation channel that is also present in our model. In particular, the response is positive and statistically significant from zero for entrepreneurs at the median of the return distribution, depicted by the dashed line. In terms of magnitudes the estimates indicate that in response to the cut in the interest rate, the median entrepreneur increases her exposure to the firm by one to two percent in the first two years after the shock. While the blue solid line shows that the response after one year is similar to the one on impact, the grey line indicates that after two years the response is a bit smaller.

Turning to heterogeneity, there is suggestive evidence that entrepreneurs at both extremes of the return distribution react relatively strongly on impact as well as one year after the shock. Through the lens of our model, and in accordance with Figure 3.6, this could be interpreted as strong direct effects for entrepreneurs with large firms and hence small returns, and large responses of entrepreneurs with small firms and hence large returns. This u-shape, however, disappears in the second year after the shock. While our model does not imply a decline in the portfolio share after two years for firm owners with low returns, as is indicated by the grey line, we would indeed expect the reallocation towards firm capital to be strongest right after the shock materializes, i.e. on impact and one year after the shock.

Discussion and robustness Given that the SCF is a repeated cross-section and not a panel, the identifying assumption that we make is that the characteristics of entrepreneurial households within a given percentile of the business return distribution stay unchanged over time. This assumption appears reasonable for entrepreneurs with firms in the middle of the firm size distribution and hence with close to median returns. Their business values and net worth by construction assume relatively common values, and therefore the households themselves are not likely to be unusual in terms of observed and unobserved characteristics. Also, the non-parametric estimation of portfolio shares ensures that for the percentiles in the middle of the return distribution we include information on portfolio shares from many neighboring percentiles. This makes the estimates for the middle percentiles less sensitive to individual observations in percentile p at year t. Therefore, we are confident that the portfolio reaction at the median of the return distribution is well identified.

Both for very small and very large firm owners, the identifying assumption might be less credible. On the one hand, there could be considerable turnover due to entry and exit of firms among the small firm owners, confounding our results at the upper end of the return distribution. Very wealthy entrepreneurs, on the other hand, with businesses producing returns at the low end of the return distribution, might possess some peculiar characteristics that are not shared by entrepreneurs at the same extreme position in the return distribution in a different year. For both of these groups, at the high and low end of the return distribution, by definition there are few neighboring percentiles either to the right or to the left of the given percentile p, and hence individual observations can influence the estimate of the portfolio share relatively strongly.

To address this issue, we first run regressions of the observed portfolio shares on various observable characteristics of the households. The controls are the same as in columns three and four of Table 3.16 in Appendix B. We then subtract the predicted portfolio shares from the actual shares, and then estimate the portfolio shares $\gamma_{p,t}$ for each year and percentile of the return distribution as described above, but this time using the residual portfolio shares. Last, we run (3.8) using the new estimates of the portfolio shares $\gamma_{p,t}$. The results are displayed in the right panel of Figure 3.8. They are very similar to the ones in our baseline specification, though the portfolio responses at the median are more muted. The general shape, however, and therefore the results regarding heterogeneity of responses highlighted above, are not affected by using residual portfolio shares instead of the actual ones.

7 Consequences of Higher Wealth Inequality

The wealth distribution is endogenous in our model. To study how higher wealth inequality affects the transmission of monetary policy, in principle we need to take a stance on what has been the exogenous force that has raised wealth inequality to a new, higher level.

Before doing so, however, in Subsection 7.1 we follow a different approach. In particular, we ask how changing the distribution affects the transmission of a cut in the interest rate, keeping all policy functions fixed at their steady state values. This allows us to stay silent on the driver of increasing wealth inequality in the recent decades. The results therefore apply to the effects of higher wealth inequality on monetary policy transmission regardless of the concrete source of elevated inequality. The disadvantage of the approach is that we can only study a static approximation to the general equilibrium response which lacks important dynamic elements, as we will discuss below.

In Subsection 7.2 we re-parameterize our model to generate a steady state with higher wealth inequality. Motivated by the decline in estate taxation in the US in recent decades, we achieve this by assuming that entrepreneurs are born with positive (inherited) wealth. We then compare the effects of monetary policy across the two model specifications, one with low and one with high inequality.

7.1 Fixed policy functions

We begin by calculating an approximate general equilibrium response to a change in the interest rate. Let Y denote aggregate output, C aggregate consumption, I aggregate investment (both into capital of private firms $\int k_e$ and into capital of the representative firm K_p), and G government expenditures. In this section, we assume that

$$Y = C(Y, r^b) + I(Y, r^b) + G(Y, r^b) \ , \tag{3.9}$$

i.e. for simplicity we ignore parts of output here that are relatively small, in particular price, portfolio and capital adjustment as well as financial intermediation costs. Taking the total differential of (3.9), we have

$$\mathrm{d}Y = \left[\frac{\partial C}{\partial Y} + \frac{\partial I}{\partial Y} + \frac{\partial G}{\partial Y}\right] \mathrm{d}Y + \left[\frac{\partial C}{\partial r^b} + \frac{\partial I}{\partial r^b} + \frac{\partial G}{\partial r^b}\right] \mathrm{d}r^b \ .$$

Rearranging yields

$$\frac{\mathrm{d}Y}{\mathrm{d}r^b} = \frac{\frac{\partial (C+I+G)}{\partial r^b}}{1 - \frac{\partial (C+I+G)}{\partial Y}} = \frac{\frac{\partial (\int_i (c_i^* + d_i^* + f_i^*)\mu(i) \ \mathrm{d}i \ + G)}{\partial r^b}}{1 - \frac{\partial (\int_i (c_i^* + d_i^* + f_i^*)\mu(i) \ \mathrm{d}i \ + G)}{\partial Y}} \ , \tag{3.10}$$

where the last fraction expresses the aggregate effect in terms of individual households' optimal decision rules. The term on the right-hand-side of (3.10) has an intuitive interpretation. The general equilibrium change in aggregate output upon a change in the interest rate is the direct effect of the interest rate change (the numerator) divided by one minus the marginal propensity to spend an additional dollar on consumption, investment or government spending (the denominator). The goal of this section is to evaluate (3.10) once with the true steady state distribution, μ^{steady} and once with a distribution that features higher wealth inequality, μ^{high}, holding the policy functions fixed.

Importantly, this approximation is static, i.e. only contemporaneous variables enter. Dynamic effects on today's household decisions, e.g. higher consumption today caused by looser borrowing constraints in the future due to elevated labor income, are shut off. We still find the approximation useful as a starting point of our analysis.

Table 3.9: Response to persistent decrease in the liquid interest rate r^b under steady state wealth distribution and distribution with higher inequality.

	Top 10% wealth share	$\frac{\mathrm{d}Y}{Y}$	$\frac{\partial(C+I+G)}{\partial r^b}\mathrm{d}r^b$	$\frac{\partial(C+I+G)}{\partial Y}$	$\frac{\partial I_e}{\partial r^b}\frac{\mathrm{d}r^b}{I_e}$
Steady state	81.5%	1.09%	.0102	.5024	2.49%
High inequality	82.5%	1.17%	.0116	.4978	2.84%
Relative change	1%	7%	13%	-1%	14%

To calculate the denominator of (3.10), we need to make an assumption on how individual household income, which we denote by y_i, fluctuates with aggregate income. To see this, rewrite

$$\frac{\partial c_i}{\partial Y} = \frac{\partial c_i}{\partial y_i}\frac{\partial y_i}{\partial Y}$$

and analogously for investment. We assume that $\forall i$, $\frac{\partial y_i}{\partial Y} = \frac{y_i}{Y}$, i.e. individual income fluctuates proportionally to aggregate income for all households.[33]

Next, we approximate the derivatives $\frac{\partial c_i}{\partial y_i}$ (MPC), $\frac{\partial f_i}{\partial y_i}$ (MPI in private firm capital) and $\frac{\partial d_i}{\partial y_i}$ (MPI in representative firm capital) as the fraction of a 500$ transfer into the liquid account that households would spend on c, f and d respectively within one quarter.[34] To calculate $\frac{\partial G}{\partial Y}$, we assume that households finance government expenditures in proportion to their income. Given the proportional tax on labor income τ_l, this is a good approximation. Formally, we assume that $\frac{\partial G}{\partial Y} = \frac{\partial\left(\int_i g_i \,\mathrm{d}i\right)}{\partial Y}$ and that $\frac{\partial g_i}{\partial Y} = \frac{\partial g_i}{\partial y_i}\frac{\partial y_i}{\partial Y} = \frac{g_i}{y_i}\frac{y_i}{Y}$, where g_i corresponds to government spending financed by household i, i.e. her taxes paid minus transfers received.

Turning to the numerator of (3.10) we require knowledge of how households react to a change in the interest rate r^b. For this, we use the results from the direct effect to the monetary policy shock in Section 5, which we plotted in Figure 3.6 for the case of private firm investment. We obtained these results from exposing households to an isolated change in the interest rate (the one that is depicted in Figure 3.5), keeping all other aggregate variables at their steady state levels. We use the impact responses of household investment and consumption respectively. From the government budget constraint we further know that $\frac{\partial G}{\partial r^b} = -B^S$.

The first row in Table 3.9 shows the results of the approximation (3.10) using the steady state distribution of households. The cut in the interest rate leads to an increase in output of 1.09%, which is a bit smaller than the response in Section 5 of 1.39%. This is not surprising, given that in the approximation we implicitly turned off all *dynamic* general equilibrium forces affecting the households' decision. In a full dynamic general equilibrium setting households realize that economic activity is stimulated and hence income is higher for several quarters. This moves households further away from the borrowing constraints,

[33] We therefore abstract from unequal incidence of aggregate income movements, in contrast, for instance, to Patterson (2019). In the next section this assumption is relaxed, as disposable income will be varying differently for different households following a monetary policy shock.

[34] In doing so, we use the same approach that KMV propose to calculate MPCs in their model.

leading to less precautionary saving, which in turn results in even more aggregate demand (see Auclert, Rognlie, et al. (2018) for a formal discussion of this argument). Our static approximation abstracts from this dynamic effect.

To construct a high-inequality wealth distribution μ^{high} we proceed as follows. Guided by the empirical evidence in Section 2, we want to increase the wealth of the average entrepreneur relative to the average worker, and in doing so generate higher wealth inequality. To focus on the role of entrepreneurs most clearly, we leave the distribution of wealth among workers μ_w unaltered, and only change the distribution of wealth for entrepreneurs μ_e. In particular, we multiply the size of each entrepreneur's private firm by a factor of $1 + x$. We choose $x = 27\%$, which generates an increase in the overall top 10% wealth share of one percentage point. This is a relatively mild increase in wealth inequality. Figure 3.15 in the appendix shows that between the early 1980s and today the top 10% wealth share has gone up by about ten percentage points. However, assuming the supply of labor in the economy remains unchanged, the employment share in the private firms significantly increases, from 41% in the steady state distribution to 46% in the new distribution which features larger private firms. We therefore view the considered experiment as a reasonable representation of the shift of wealth from workers to entrepreneurs that actually occurred in the US between 1980 and today.

The core result of this section is that once we move to the distribution that features higher wealth inequality, the output response is strengthened, in our experiment by 0.08 percentage points (second row of Table 3.9). This represents an increase of the aggregate output response of 7% compared to the response implied by the original steady state distribution. Strikingly, as the relative changes in the last row reveal, the entire increase of the effects on real activity stem from the direct effect of monetary policy, i.e. from a change in the numerator of (3.10). The average marginal propensity to consume and invest in the economy (second to last column) is almost unaltered by changing the distribution.[35]

The larger effects under higher inequality can be explained with the mechanisms discussed in Section 5. A shift towards a distribution that features higher wealth inequality puts more wealth into the hands of entrepreneurs with a strong portfolio reallocation response. This can be seen in Figure 3.9, where the orange dashed line shows the direct effect of the change in the interest rate $\{r_t^b\}$ on firm investment (this is the same as in Figure 3.6). Under the new distribution with higher inequality, as the distribution of firm capital is shifted to the right, there exist more entrepreneurs owning large firms. Since these entrepreneurs exhibit a large elasticity of firm investment, monetary policy has stronger effects. In sum, the direct effect of the shock on aggregate private firm investment rises by 14% when moving from the

[35] There are two counteracting effects on the denominator of (3.10). Higher inequality leads to smaller aggregate marginal propensities to invest because of a *within*-occupation effect. Small firm owners have higher propensities to invest additional income into their firms as large firm owners, and when more capital is in the hand of the large firm owners, overall MPIs decline. Higher inequality leads to larger aggregate marginal propensities to invest, however, because of a *between*-occupation effect. The average entrepreneur features a higher MPI than the average worker, and in our experiment we increase the share of wealth held by the entrepreneurs. These two effects approximately cancel out in our calibration.

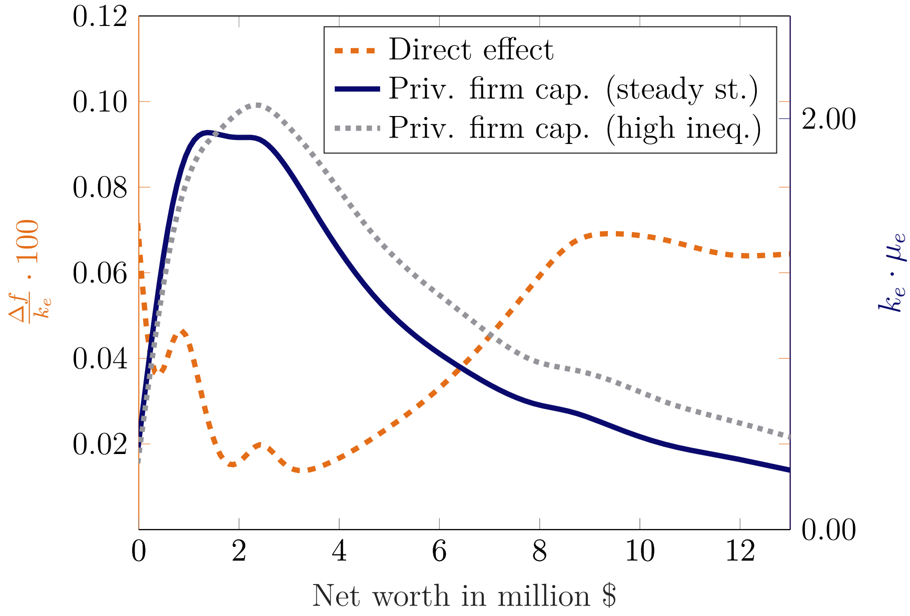


Figure 3.9: Direct effect and distributions of firm capital under low and high inequality.
Notes: Orange dashed line (left y-axis): Change in firm investment caused by direct effect, i.e. only by change in liquid interest rate r^b, other prices fixed. Blue solid line (right y-axis): Share of private firm capital $k_e \cdot \mu_e$ in steady state. Grey dotted line (right y-axis): Share of private firm capital $k_e \cdot \mu_e$ under counterfactual distribution with higher wealth inequality.

low- to the high-inequality distribution, as the last column of Table 3.9 shows.

One caveat of the results so far is that wealth inequality within the group of entrepreneurs *declines* in our experiment. For instance, the wealth held by the richest ten percent of entrepreneurs relative to the wealth held by all entrepreneurs decreases by about 2.5 percentage points. This stands in contrast to our empirical findings in Section 2, which indicate that inequality among entrepreneurs has actually risen since the 1980s. To account for this, we consider a second high-inequality distribution μ^{high}. This time, instead of multiplying the firm size of all entrepreneurs by $1 + x$ we only increase the firms of the top 10% richest entrepreneurs, more precisely of those whose firms are among the 10% largest. We choose $x = 47\%$, again to generate an increase in the top 10% wealth share across all occupations of one percentage point. In contrast to before, such a shift of wealth leads to an increase in inequality even among entrepreneurs: the top 10% share among entrepreneurs rises by about 2.5 percentage points. The output response in this case is 10% larger compared to the baseline scenario with μ^{steady}. The relatively stronger amplification compared to that shown in Table 3.9 is intuitive, as now the shift of wealth towards the (high-elasticity) wealthy entrepreneurs is even more pronounced than it was in the first experiment.

7.2 General equilibrium response under higher wealth inequality

We now turn to the full dynamic general equilibrium analysis which shows that the results obtained from the approximation in the previous subsection carry over when taking into account dynamic effects. In principle, there are many ways to endogenously create more inequality in the steady state of our model. Here, we opt for one that leads only to a small deviation from the original model. This makes it possible to most clearly attribute the differential responses we find under the new specification to higher wealth inequality.[36] In particular, we assume that entrepreneurs are born with a positive amount of assets, unlike in the initial version of our model, in which we assumed that all households were born with zero assets. We find that this small change to the model is enough to endogenously generate a higher degree of wealth inequality in steady state.

One can interpret this modification as decreasing progressivity of the US tax system, well-known to have taken place since the 1980s (see, for instance, Hubmer et al. (2021)). More specifically, our experiment can be interpreted as a decrease in estate taxation, leading to higher initial endowments of wealthy households' heirs. Federal estate taxation has become less broad-based since the 1980s. For instance, while in 1982 2.8% of the US adult population had assets exceeding the threshold that made them subject to estate taxation upon death (assets greater than 325,000$, or 808,000$ in 2016 terms), this fraction decreased to only 0.32% in 2016 (assets greater than 5,450,000$). At the same time, the top bracket tax rate applied to these estates has declined, from 70% in the early 1980s to 40% today.[37]

We make the following adjustment to our quantitative model to implement this modification. Instead of assuming that entrepreneurs are born with zero wealth, we now assume that they own positive wealth $w' = \{b', a', {k_e}'\}$ right from the start of their lives. We only allow entrepreneurs, not workers, to be born with positive wealth in order to identify most clearly the effect of the shift in wealth towards the private entrepreneurs that we observe in the data. We also refrain for simplicity from making any adjustment to the preferences of the households, implicitly assuming that they do not derive utility from bequeathing wealth to their offspring. The difference between the exercise conducted here and the one in the previous section is that this time the steady state distribution (and with it all prices) endogenously adjust in response to the new environment.

To make our results here comparable to the exercise in Section 7.1, we calibrate the amount of inherited wealth w' such that, as above, we obtain an increase in the top 10% wealth share of one percentage point in the new steady state. It turns out that we require entrepreneurs to be born with about \$590,000 to achieve this.[38]

We now expose the re-parameterized model to the same monetary policy shock as before.

[36] Another approach would be to re-calibrate the model, targeting for instance a higher employment share in private firms. Then, however, all calibrated parameters would change, and it would be less clear which changes are driving our results.

[37] These numbers are taken from information provided by the Internal Revenue Service, in particular from SOI Bulletin articles by Barnes (2021), Jacobson et al. (2007), and Schwartz (1988), as well as from information provided on the IRS website, accessed on Nov 29, 2020, at https://www.irs.gov/businesses/small-businesses-self-employed/whats-new-estate-and-gift-tax.

[38] We assume that the endowed wealth w' is composed entirely of private firm capital, i.e. $b' = a' = 0\$$. We found this assumption to be inconsequential for the main results of this section. In order to account for the non-zero bequests, we make the corresponding adjustments to the interest rates in our model. In particular, we assume that the interest rates on all three assets are proportionally adjusted downward by the same factor in order to finance the endowments of newborn entrepreneurs.

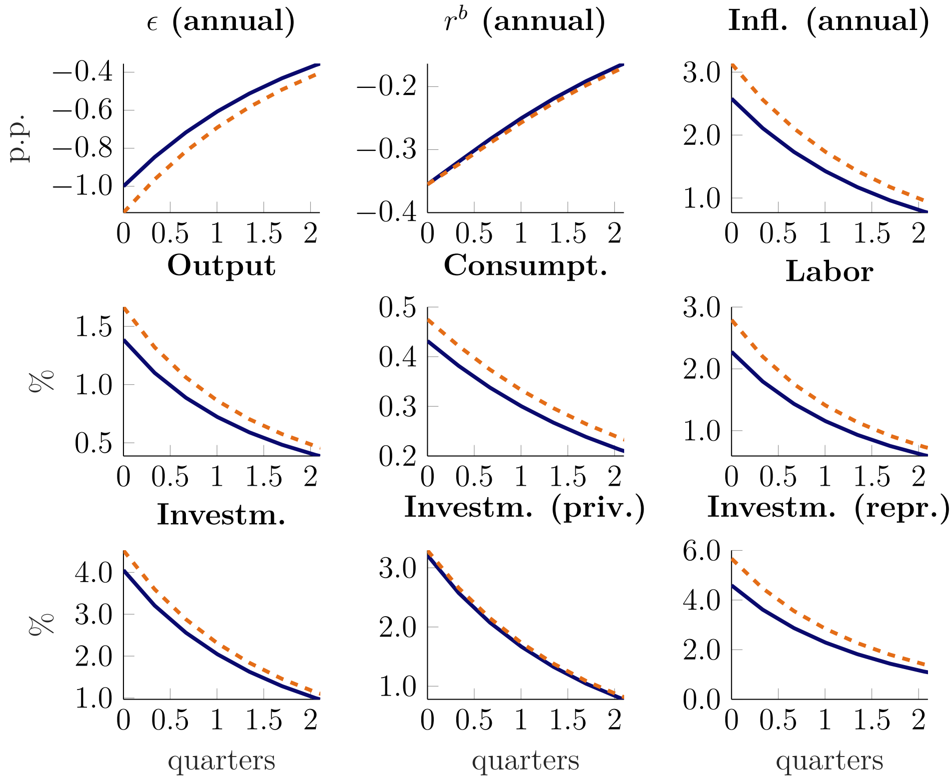


Figure 3.10: Response of aggregate variables to an expansionary monetary policy shock under low and high wealth inequality.
Notes: Depicted are deviations from steady state values. The solid blue line shows the response to the shock under the original steady state distribution. The dashed orange line shows the response when newborn entrepreneurs are born with positive wealth w', which leads to higher wealth inequality in steady state. All high-inequality responses are rescaled to imply the same impact drop of r^b as under low inequality.

To be able to compare the responses under the two scenarios, we rescale them to imply the same peak response of the interest rate r^b. Figure 3.10 plots the IRFs both for the economy with low and with high inequality. In order to make the differences between the two economies clearly visible the figure only displays the responses for the first two quarters after the shock.

As in Section 7.1, we find that the real effects of monetary policy are amplified by higher wealth inequality. The output response goes up from 1.39% under low inequality to 1.67% under high inequality, i.e. we find that the effects are magnified by about 20% of the initial response. This corroborates our previous finding, namely that increasing wealth inequality, by putting more wealth into the hands of wealthy entrepreneurs, strengthens the effects of monetary policy in general equilibrium.

At first sight, Figure 3.10 seems to indicate that private firm investment is not the most influential driver behind the stronger output effects, as its response under high and low inequality are similar. To better understand the reasons for this, Figure 3.11 plots the firm capital share distribution in the initial steady state (blue solid) and its counterpart when entrepreneurs are born with positive bequests (grey dotted), as well as the direct effect on firm investment in the initial steady state. While the considered experiment does

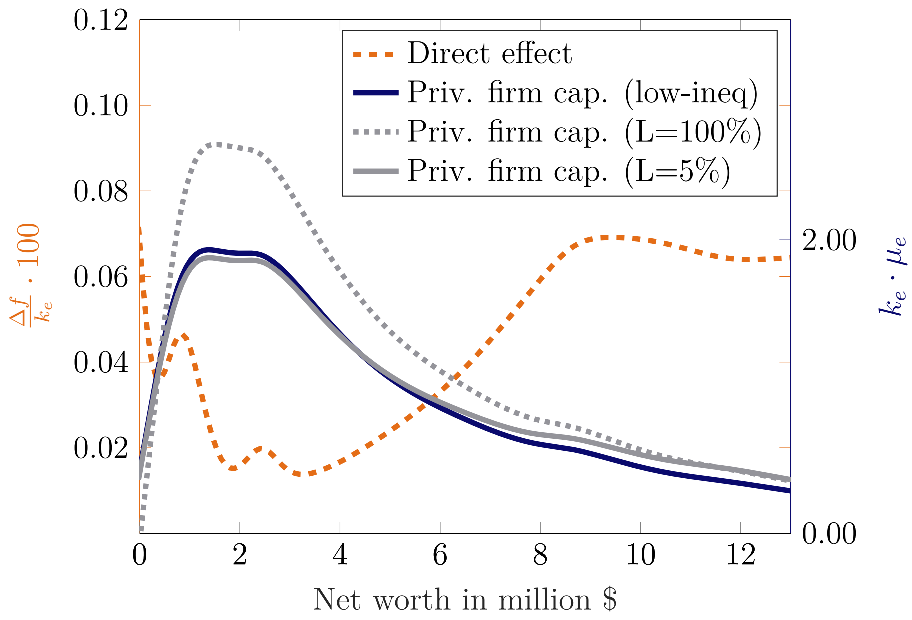


Figure 3.11: Direct effect and different private firm capital distributions.

Notes: Left y-axis: Change in firm investment caused by direct effect, i.e. only by change in r^b (evaluated in the initial steady state) (orange dashed). Right y-axis: Share of private firm capital $k_e \cdot \mu_e$. Blue solid: initial steady state distribution; grey dotted: high-inequality distribution when all entrepreneurs receive bequests; grey solid: high-inequality distribution when $L = 5\%$ of entrepreneurs receive bequests.

increase the firm capital of rich entrepreneurs with large direct effects, it increases to an even bigger extent the amount of firm capital held by medium-wealthy entrepreneurs with small investment elasticities.

Why do we still obtain larger responses of aggregate investment and output under high than under low inequality? To develop an answer, consider Table 3.10. The first row captures the initial (low-inequality) steady state of our model. Here, entrepreneurs do not receive bequests and the share of workers employed in the private firms is 41%. The second row corresponds to the high inequality steady state, in which all newborn entrepreneurs receive bequests of $590,000. The share of labor employed in the private firms rises to 48% and the initial output response to the monetary policy shock gets amplified by 20% as shown in Figure 3.10.

The last three columns split the direct effect of the interest rate change on impact by type of investment: total investment, private firm investment and representative firm investment. To obtain these numbers we compute the reaction of all households in the economy to the path of the interest rate $\{r_t^b\}$ depicted in Figure 3.5, exactly like we did before in Section 5. The direct effect on private firm investment (second to last column) is indeed smaller under high inequality (second row) than under low inequality (first row), just like Figure 3.11 suggested. Total investment still responds more strongly, however, since in the new steady state a larger share of it goes into private firm capital. Since private firm investment features a higher elasticity than investment into the representative firm, the direct effect on total investment becomes larger. This in turn causes incomes to rise more strongly under

Table 3.10: Responses to interest rate changes under different bequest regimes.

(1) L	(2) Bequest size (M$)	(3) Y response increase	(4) Labor at entre.	(5) Top 10% wl. share	(6) Direct effect on Inv. $\frac{\partial I}{\partial r^b}\frac{\mathrm{d}r^b}{I}$	(7) $\frac{\partial I_e}{\partial r^b}\frac{\mathrm{d}r^b}{I_e}$	(8) $\frac{\partial I_r}{\partial r^b}\frac{\mathrm{d}r^b}{I_r}$
0%	-	-	41%	81.5%	1.42%	2.49%	0.73%
100%	0.59	20%	48%	82.5%	1.53%	2.40%	0.72%
50%	1.18	12%	46%	82.9%	1.48%	2.37%	0.71%
10%	5.89	8%	43%	83.3%	1.48%	2.52%	0.71%
5%	11.79	5%	42%	83.1%	1.47%	2.53%	0.72%

Notes: The first row represents the initial steady state, the second to last rows vary the bequests received by entrepreneurs. Column (1) shows the fraction of entrepreneurs who receive bequests. Column (2) shows the bequeathed amount per bequest in million \$. Column (3) shows by how much the impact output response is increased compared to the baseline. Column (4) indicates the fraction of workers employed in the private firms in steady state, column (5) the top 10% wealth share. The last three columns correspond to the impact change in total investment (6), in private firm investment (7) and in investment into the representative firm (8) that is due to the direct effect, i.e. when all prices are at their steady state level and only $\{r_t^b\}$ evolves as in Figure 3.5.

high inequality, thereby inducing second-round (indirect) effects that in sum lead to the larger output response shown in Figure 3.10.

The result that the direct effect on private firm investment *decreases* under high inequality, contrary to the experiment in Subsection 7.1, is however sensitive to the precise implementation of the change in estate taxation. Suppose only a "lucky" fraction L of entrepreneurs receives bequests, and the fraction $1 - L$ is born with zero assets as in the initial steady state. We continue to assume that the total amount of bequests is the same as before, \$590,000 per entrepreneur. Hence, when the share of lucky entrepreneurs is higher, bequests are more concentrated. Fewer entrepreneurs receive bequests, but for those who do the bequest is larger.

A graphical illustration of the steady state firm capital distribution with $L = 5\%$ is depicted in Figure 3.11 (grey solid line). The difference compared to the initial distribution in this case is smaller than when $L = 100\%$. The reason is that when $L = 5\%$ the few entrepreneurs who are lucky recipients of bequests diversify the substantial wealth they inherit across the three assets relatively quickly, and therefore the increase in aggregate private firm capital compared to the baseline is limited. In contrast, when $L = 100\%$ the individual bequests are much smaller and because returns on private firm capital are high for small firms, newborn entrepreneurs leave the inherited wealth predominantly in firm capital. Figure 3.11 also shows that, in contrast to the case where we assumed bequests for all entrepreneurs, when $L = 5\%$ only the right tail of the firm capital distribution becomes notably fatter, indicating the presence of a larger share of entrepreneurs with a strong direct investment response to interest rate changes.

The last three columns of Table 3.10 echo these observations. The fraction of workers employed in the private firms falls as L decreases, because the higher the bequests the more newborn entrepreneurs invest the inherited wealth in a diversified portfolio. The direct effect of private firm investment increases relative to the initial steady state when only a small fraction of entrepreneurs is born wealthy, because the share of high-elasticity entrepreneurs grows.[39] In sum, the output response is amplified in all experiments under higher inequality. When rescaled to imply an increase in the overall top 10% wealth share of one percentage point, the magnitude of amplification varies between 3% and 20%.

8 Conclusion

Entrepreneurs constitute a small fraction of all households, but they hold a large share of total wealth, and their firms employ a larger share of the workforce. In addition, the gap between average wealth held by entrepreneurs and by workers has increased over the recent decades. We highlighted these facts using survey data from the SCF. Together with a well-documented rise in wealth inequality in the US since the 1980s, these observations motivated our research questions: How important are entrepreneurs for the transmission of monetary policy to the real economy? How does the observed shift in wealth towards entrepreneurs affect the transmission of monetary policy?

We built a HANK model with entrepreneurs to provide answers to these questions. Upon a cut in the interest rate on liquid assets, entrepreneurs increase the portfolio share of their private firms. The strength of this portfolio reallocation effect is heterogeneous across the net worth distribution. Wealthy entrepreneurs own large firms and earn business returns that are only slightly above the risk-free rate, both in our model and in the data. When monetary policy changes the risk-free rate, their excess return is therefore affected significantly in relative terms. Hence, they rebalance their portfolio more strongly than entrepreneurs with low net worth, whose excess returns are high and therefore affected relatively little by a change in the interest rate. We presented empirical evidence that two key implications of our model, decreasing business returns in net worth and a (heterogeneous) portfolio reallocation effect, are supported by the data.

In our model, entrepreneurs are quantitatively important for the impact of monetary policy on the real economy. If entrepreneurs do not respond to changes in prices and aggregate quantities, the output response to an interest rate cut is approximately 50% smaller. Moreover, our model implies that an increase in the top 10% wealth share by one percentage point—a mild increase in wealth inequality—amplifies the aggregate output response by 3 to 20%.

There are additional aspects of entrepreneurial investment that could be important for the

[39] The former case corresponds relatively better to the empirical observation that wealth inequality among entrepreneurs has increased over time. While the experiments in this subsection do not account for this additional feature of the data, the top 10% share within the group of entrepreneurs in fact decreases compared to the initial steady state when all newborn entrepreneurs receive bequests. In contrast, in the case in which only very few entrepreneurs receive bequests, inequality among them stays nearly unchanged. Hence, in this regard the latter case is relatively more in line with the empirical evidence than the former.

transmission of monetary policy that we have abstracted from. We did not model entry and exit from entrepreneurship, that might in addition vary over the business cycle (see Levine and Rubinstein, 2020, and references therein). We also did not allow for collateralized borrowing or financing through outside equity. We partly did so because a predominant fraction of entrepreneurs relies on their own funds to finance their firm (Table 3.14). Lastly, we ruled out the possibility of running multiple businesses at a time, which could have implications for the level and variance of the return that firm owners obtain from their overall business wealth. We are confident that the portfolio reallocation mechanism we highlighted continues to be important in a model capturing these additional features, but we leave these issues for future work.

Our framework can be used in future research to understand how specific shocks to the financial situation of private business owners affect monetary policy transmission. If entrepreneurs are in a dire financial situation, lacking the resources to raise investment into their firm, a crucial link in the transmission chain is missing and monetary policy might become less powerful. This force was likely at play during the pandemic-induced lockdowns in the previous two years, which severely affected business revenues (Fairlie, 2020). If monetary policy is to function as usual, with entrepreneurs transmitting interest rate cuts to the economy by raising investment, this provides a rationale for government programs helping the affected firms and their owners (e.g. the Paycheck Protection Program, see Granja et al., 2020; Hubbard and Strain, 2020).

Appendices

A Empirical Evidence on Entrepreneurs and Wealth Inequality

Aggregate statistics This subsection provides additional empirical evidence on entrepreneurial households and their firms. Figure 3.12 plots the share of households that we classify as entrepreneurs over time. Figures 3.2 and 3.14 in the main text documented the ratio of average wealth held by entrepreneurs and non-entrepreneurs, as well as average employment per firm respectively. Figures 3.13 and 3.3 display the same information, only this time in terms of aggregate statistics. Figure 3.13 shows that over the recent decades, on average a third of wealth was held by the entrepreneurs in the economy. Furthermore, the same upward time trend that was visible in the plots of the main text is visible here as well. As was the case for the average wealth ratios in the main text, one might be worried that this result is driven exclusively by the data point in 1983 which displays a very low entrepreneurial wealth share. To alleviate this worry, we again conduct a t-test asking whether the linear time trend is statistically different from zero. We find that even when excluding the observation in 1983, the time trend is still positive and statistically significant with a p-value of 1.5%.

Figure 3.3 in the main text documented the share of employment in the entrepreneurial firms of all employment in the US. The share is large (on average 46%) and growing over time. We already mentioned in the main text that we lack information on the intensive margin of labor supplied in the entrepreneurial firms. One might therefore be worried that we overestimate the share of employment in the entrepreneurial firms if private firms were systematically more likely to employ people on a part-time basis. However, we believe that this should not be a major problem here. According to the Bureau of Labor Statistics (BLS), in January 2019, about 17% of all US workers were considered to be working part-time, the large majority of them (13%) for non-economic reasons. Of the wage and salary workers among those who worked part-time for non-economic reasons more than 50% worked in industries such as retail trade, food services and drinking places, and private educational services.[40] While data on industry affiliation of our entrepreneurs' firms in the more recent waves of the SCF is too coarse to be informative, data on industries up until 1992 is fine

[40] https://www.bls.gov/opub/mlr/2018/article/who-chooses-part-time-work-and-why.htm , accessed on Nov 7, 2021.

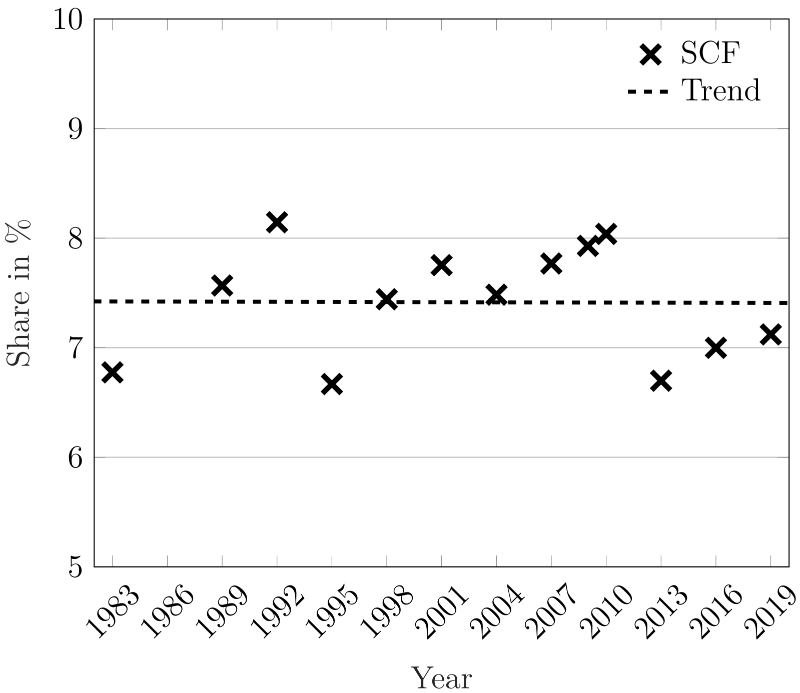


Figure 3.12: Entrepreneurs as a share of all households.

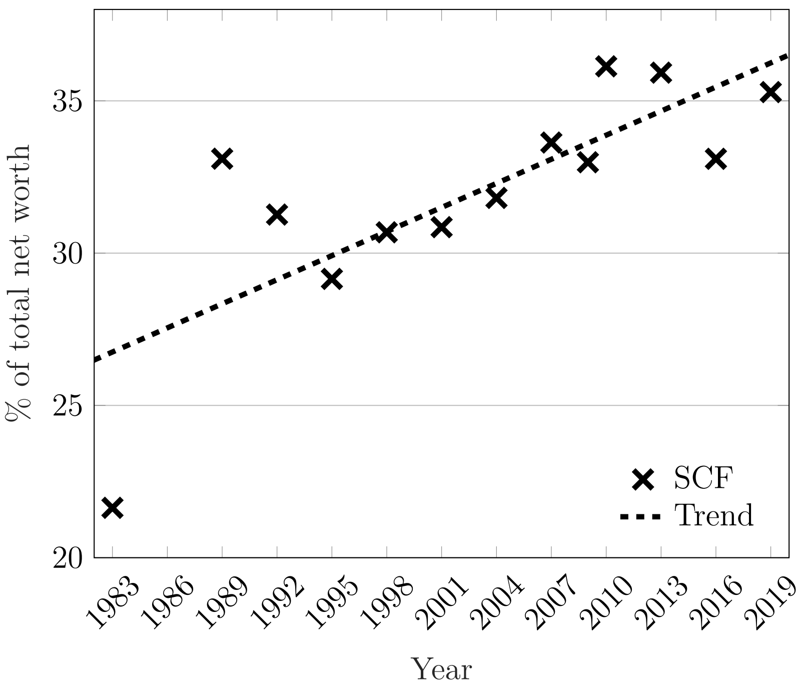


Figure 3.13: Share of total US wealth held by entrepreneurs.

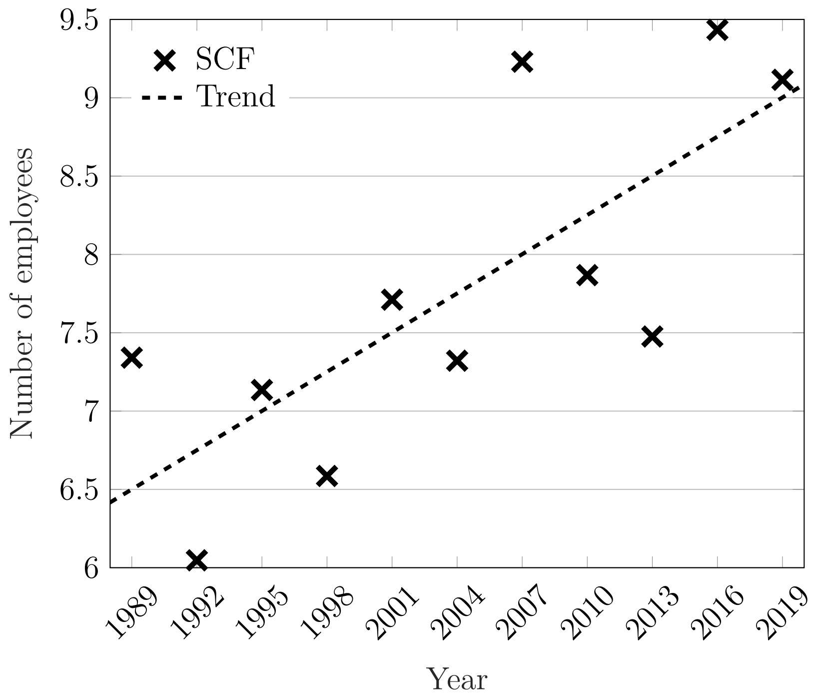


Figure 3.14: Average number of employees in firms owned by entrepreneurs.

enough to gain some insight. In 1992, the four industries most often mentioned as the industry of their first firm by our entrepreneurs were (see also Table 3.15)

- Professional practice, incl. law, medicine, architecture; accounting; bookkeeping (17%)
- Contracting; construction services; plastering; painting; plumbing (14%)
- Retail and/or wholesale business excluding restaurants, bars, direct sales (e.g. Tupperware), gas stations, and food and liquor stores (10%)
- Farm; nursery; train dogs; forest management; agricultural services; landscaping; fisheries (9%)

Hence we conclude that the overlap between those industries that feature lots of entrepreneurial activity and those in which much part-time work is reported does not appear to be a major concern.

Firm size distribution Table 3.11 documents the firm size distribution in 2019 by employment. The distribution is highly skewed: A large fraction of firms is very small in terms of employment, and only few are very large. However, as can be seen in the second column of the table, the few large firms are quite important when accounting for the overall employment in the entrepreneurial firms. More than three quarters of all employment in the private businesses is due to firms that employ ten or more workers. A similar pattern emerges from Table 3.12, which reports the firm size distribution in terms of gross sales. Lastly, Tables 3.12, 3.14 and 3.15 show the distribution of entrepreneurial firms across different legal statuses, sources of funding and industry respectively.

Table 3.11: Firm size distribution by employment (SCF 2019)

Employees	Share of firms (in %)	Share of employment (in %)
1 (Micro)	39.3	4.7
2–9 (Micro)	50.2	20.0
10–49 (Small)	7.9	17.9
50–249 (Med.)	2.2	24.2
250 and more	0.4	33.2

Notes: Since we observe employment and ownership shares only in the first two businesses of a given entrepreneurial household, we assume that if an entrepreneur has more than two businesses employment in these additional businesses is as in the second business.

Table 3.12: Firm size distribution by gross sales (SCF 2019)

Sales	Share of firms (in %)	Share of sales (in %)
< 2 Mio \$ (Micro)	93.0	14.2
2–10 Mio \$ (Small)	5.1	11.8
10–50 Mio \$ (Med.)	1.4	14.7
50 Mio \$ and more	0.5	59.3

Notes: Since we observe gross sales and ownership shares only in the first two businesses of a given entrepreneurial household, we assume that if an entrepreneur has more than two businesses gross sales in these additional businesses is as in the second business.

Table 3.13: Firms by legal status (SCF 2019)

Legal status	Share (in %)	Share of net worth (in %)
Partnership	5.9	6.6
Sole Proprietorship	40.8	13.5
S Corp.	14.1	25.8
Other Corp. (incl. C Corp.)	7.1	9.8
Limited Partnership / LLP	32.1	44.3

Notes: Left column: Share of entrepreneurs who declare that their first business is of a given legal status. Right column: Net worth of entrepreneurs who declare that their first business is of a given legal status, relative to total entrepreneurial net worth.

Table 3.14: Firms by source of funding (SCF 2019)

Source of funding	Share (in %)	Share of net worth (in %)
Personal savings	26.3	14.6
Credit card	8.5	6.3
Personal loan	6.7	6.7
Business loan	10.7	16.2
Equity investors	1.1	2.0
Inherited	0.0	0.0
No external money	53.4	54.7
No answer	1.7	9.0

Notes: Multiple answers possible. Left column: Share of entrepreneurs who declare that they used a given source of funding for their first business. Right column: Net worth of entrepreneurs who declare that they used a given source of funding for their first business, relative to total entrepreneurial net worth.

Table 3.15: Firms by industry (SCF 1992, seven most important by share)

Industry	Share (in %)	Share of NW (in %)
Professional practice, incl. law, medicine, architecture; accounting; bookkeeping	16.8	19.7
Contracting; construction services; plastering; painting; plumbing	13.8	9.4
Other retail and/or wholesale business	9.9	10.7
Farm; nursery; train dogs; forest managem.; agricultural services; landscaping; fisheries	9.4	7.0
Real estate; insurance	7.3	15.8
Manufacturing, incl. printing/publishing	6.6	11.1
Personal services: hotel, dry cleaners, funeral home	6.3	2.9

Notes: Left column: Share of entrepreneurs who declare that their first business is in a given industry. Right column: Net worth of entrepreneurs who declare that their first business is in a given industry, relative to total entrepreneurial net worth.

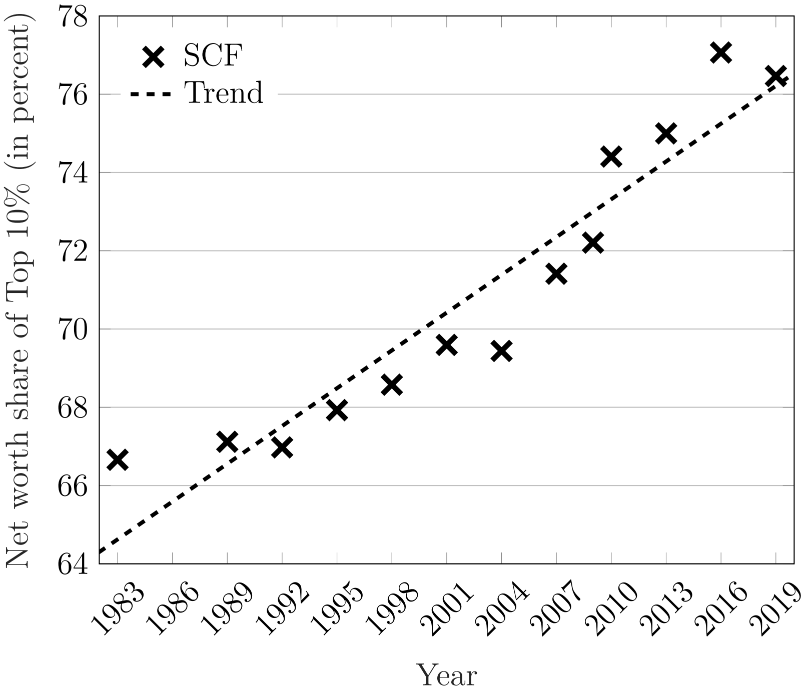


Figure 3.15: Share of total wealth held by the Top 10% wealthiest households.

Wealth inequality As pointed out in the main text, several authors have found that wealth inequality has been rising in the US since the 1980s (Hubmer et al., 2021; Kuhn et al., 2020; Saez and Zucman, 2016). Figure 3.15 shows this using the share of wealth held by the richest 10% of the population as a measure of wealth inequality. Moreover, even when conditioning on the group of entrepreneurs inequality has increased, as Figure 3.16 shows.

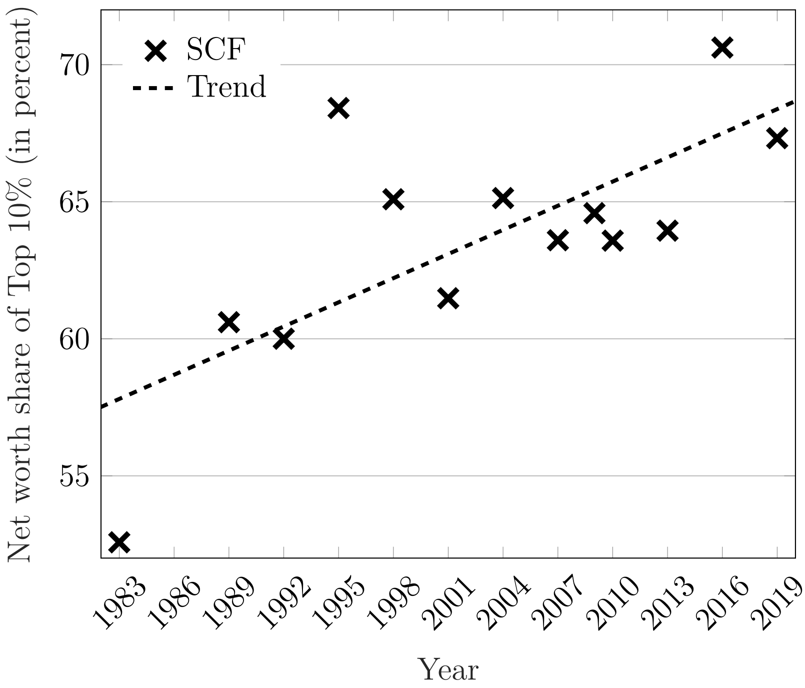


Figure 3.16: Share of total wealth held by the Top 10% wealthiest households, conditional on being an entrepreneur.

B Further Empirical Evidence on Business Returns

First, we estimate the relationship between net worth and business returns non-parametrically using kernel-weighted local polynomial smoothing with an Epanechnikov kernel. Figure 3.17 shows the results, which paint a very similar picture as Figure 3.7 in the main text. Households with net worth of \$100,000 make an average return of almost 100% while households with net worth around \$50,000,000 are estimated to earn an average return of about 35%. Note that we are considering average returns here which are larger than the median returns in Figure 3.7 as the distribution of returns is right-skewed.

Next, we estimate the relationship between net worth and business returns from the linear regression

$$r^e_{it} = \alpha + \beta \ln(\text{net worth}_{it}) + \gamma X_{it} + u_{it}. \tag{3.11}$$

Here, X is a vector of controls and u is an error term. The coefficient of interest is β which tells us by how many percentage points business returns change when the log of net worth increases by one percent.

Table 3.16 presents the estimates obtained when pooling all SCF waves. The first column shows the estimate for β we get without any additional controls. It is significantly smaller than zero and tells us that a 1% increase in net worth is associated with a decline in the return on business investment of 0.148 percentage points. In column three we control for household demographics such as age, education, marital status and the number of children. We also include fixed effects for the legal form of the business, the household's self-reported

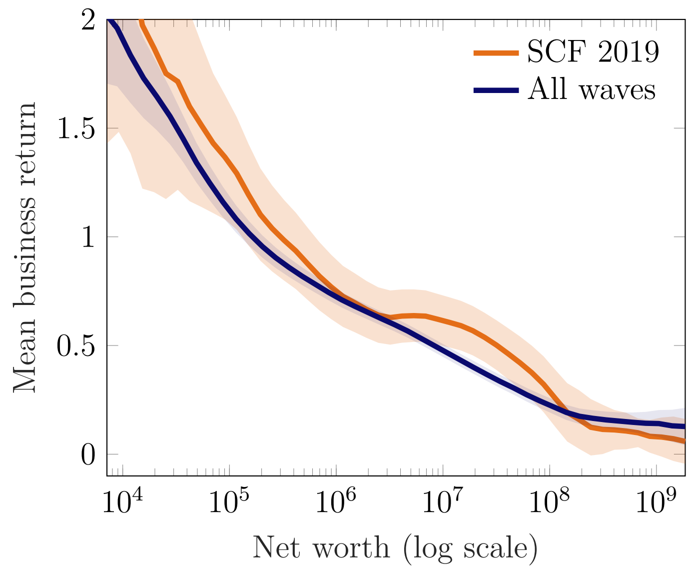


Figure 3.17: Non-parametrically estimated relationship between business returns and net worth.

Notes: We use kernel-weighted local polynomial smoothing with an Epanechnikov kernel, confidence intervals are at the 95% level.

risk attitude and the survey year. With these controls, we get an estimate for β of -0.165.

As entrepreneurs can potentially hold multiple private businesses at the same time—something that we have abstracted from in our model—we additionally control for the number of businesses the entrepreneur operates in columns three and four. We find a negative effect on the return of total business capital, which appears intuitive. By running multiple businesses entrepreneurs are able to diversify their portfolio so that the idiosyncratic risk associated with their total business investment becomes smaller. Hence, they are willing to accept a lower risk premium so that average returns on total private business investment are lower for them.

The reader might be worried because firm value affects our measure of business returns negatively as it enters the denominator. Therefore, other things equal, households that overstate the value of their business exhibit smaller business returns as well as higher net worth. As a result, measurement error in business value could mechanically lead to a negative relationship between returns and net worth.

To account for this, we replace net worth with non-business wealth in the regression equation (3.11). The estimates are shown in columns two (without controls) and four (with controls). We again find a statistically significant negative relationship, as would also be implied by our quantitative model. Households whose non-business wealth is one percent larger, earn a return on their business that is on average 0.05 percentage points lower.

The effect is somewhat smaller than the one we obtain for net worth. One reason is the possible effect of measurement error in business value mentioned above, which could bias the estimates in columns one and three. However, in an environment with decreasing returns, there is also an economic reason. If a household's net worth increases by 1%, she invests into

Table 3.16: Regressions of business returns on net worth and non-business wealth using SCF since 1989.

	(1)	(2)	(3)	(4)
Log net worth	-0.148***		-0.165***	
	(0.00455)		(0.00635)	
Log non-business wealth		-0.0795***		-0.0500***
		(0.00504)		(0.00643)
Number businesses owned			-0.0236***	-0.0463***
			(0.00327)	(0.00389)
Demographics	No	No	Yes	Yes
Legal form FE	No	No	Yes	Yes
Risk attitude	No	No	Yes	Yes
Year FE	No	No	Yes	Yes
Observations	7288	7132	7288	7132

Notes: Demographics include age, dummies for education level, number of kids, marital status, whether the entrepreneur founded the business, and the years that have passed since the start/acquisition of the business. Risk attitude is captured by a categorical variable with four categories constructed from the respondent's answer to the question: "On a scale from zero to ten, where zero is not at all willing to take risks and ten is very willing to take risks, what number would you be on the scale?"

her firm but at the same time increases the portfolio share of non-business assets because of the diminished return. Therefore, a 1% increase in net worth is associated with an increase in non-business wealth of more than 1%.

Portfolio shares Another testable implication of our model is that for owners of small firms, firm capital makes up a larger share of their portfolio. In other words, exposure to business wealth declines in net worth. We do not find robust evidence supporting this in the SCF data. When regressing portfolio shares (i.e. the empirical analogue of $k_e/(k_e + a + b)$) on the log of net worth ($k_e + a + b$), typical OLS estimates we find are close to zero, not smaller than zero as our model implies. There are two main explanations for this.

First, as mentioned above, there could be measurement error in k_e, which would tend to bias simple OLS estimates, as k_e appears on both sides of the equation. When we regress the ratio $k_e/(a+b)$ on the log of non-business wealth $(a+b)$, trying to alleviate this problem, we do find significant negative coefficients. This would imply that when non-business wealth increases, the portion of firm wealth in the household's portfolio decreases, i.e. falling exposures to firm capital, as suggested by our model. However, the issue of simultaneity clearly also affects this regression.

Second, small firm owners might want to hold other assets than their firm for objectives that we abstracted from in our model. For instance, households might want to hold some liquid assets for transactional purposes, which we do not model. Also, if we endogenized the occupational choice, households would save up assets before actually starting their firm, as in Cagetti and De Nardi (2006). In this case, at the start of their entrepreneurial career they would not be holding a portfolio that is almost completely made up of their firm, as is the case in our model.

C Model Details and Derivations

C.1 Equilibrium

An equilibrium is defined as paths for household decisions $\{a_t, k_{et}, b_t, c_t, d_t, \ell_t, f_t, n_{et}\}_{t\geq 0}$, input prices $\{r_t^k, w_t\}_{t\geq 0}$, returns on liquid and illiquid assets $\{r_t^b, r_t^a\}_{t\geq 0}$, the share price $\{q_t\}_{t\geq 0}$, the intermediate good price $\{p_t\}_{t\geq 0}$, the inflation rate $\{\pi_t\}_{t\geq 0}$, government transfers $\{T_t\}_{t\geq 0}$, distributions $\{\mu_{wt}, \mu_{et}\}_{t\geq 0}$, and aggregate quantities such that, at every t:

1. Given prices and aggregate quantities implied by the distributions $\{\mu_{wt}$ and $\mu_{et}\}_{t\geq 0}$, and the stochastic processes for individual states, policy functions c_w^*, b_w^*, a_w^*, and ℓ^* for workers, and c_e^*, b_e^*, a_e^*, k_e^*, and n_e^* for entrepreneurs solve the problems (3.2) and (3.3).

2. All firms optimize, given input prices.

3. The sequence of distributions satisfies aggregate consistency conditions

4. The bond market and the capital market clear, the labor market clears, and all goods markets clear.

5. Monetary policy follows the Taylor rule (3.7), the government budget is balanced (3.6).

Bond market clearing Households demand liquid bonds

$$B_t^D = (1 - s_e) \int b_w^*(i)\, \mu_{wt}(i)\, \mathrm{d}i + s_e \int b_e^*(i)\, \mu_{et}(i)\, \mathrm{d}i$$

and hence the bond market clears if $B^S = B_t^D$.

Capital market clearing The representative firm optimally demands K_{pt}. Total supply of the illiquid asset by the households is given by

$$A_t^S = (1 - s_e) \int a_w^*(i)\, \mu_{wt}(i)\, \mathrm{d}i + s_e \int a_e^*(i)\, \mu_{et}(i)\, \mathrm{d}i$$

and hence the capital market clears if $K_{pt} = A_t^S - q_t$.

Labor market clearing Aggregate labor supply is given by

$$N_t^S = (1 - s_e) \int \exp(z) \cdot \ell^*(i)\ \mu_{wt}(i)\ \mathrm{d}i\ .$$

Labor demand is the sum of demand by public firms and demand by entrepreneurial firms

$$N_t^D = N_{pt} + N_{et} = N_{pt} + s_e \int {n_{et}}^*(i)\ \mu_{et}(i)\ \mathrm{d}i\ .$$

and hence the labor market clears if $N_t^D = N_t^S$.

Input goods market clearing We have that

$$\int_0^1 Y_t(j)\mathrm{d}j = Y_{pt} + s_e \int y_e(i)\mu_{et}(i)\ \mathrm{d}i$$

Intermediate goods market clearing We have that

$$Y_t = \int_0^1 Y_t(j)\mathrm{d}j$$

and by symmetry of all intermediate goods firms,

$$Y_t(j) = Y_t \quad \forall j$$

Once the markets for the inputs and the intermediate goods clear, market clearing for the final good is implied by Walras' law.

Entrepreneurial taxes Revenues from taxing entrepreneurs that show up in (3.6) are defined as

$$\mathrm{Rev}_t = s_e\tau_e \int \left(\Pi_e(i) - \delta k_{et}\ \mu_{et}(i)\right) \mathrm{d}i$$

C.2 Hamilton-Jacobi-Bellman equation

If labor productivity and entrepreneurial talent follow the jump-drift processes described in Section 4, the solution to the entrepreneurs' problem can be characterized recursively by the following Hamilton-Jacobi-Bellman equation for low-productivity types (y_l)

$$\begin{aligned}(\rho+\zeta)\,V(b,a,k_e,y_l) &= \max_{c,d,f} u(c,\bar{\ell}) \\ &+ V_b(b,a,k_e,y_l)\Big[(1-\tau_e)\Pi_e(k_e,y_l) + r^b(b)b + T - d - \chi^a(d,a) \\ &\qquad - f - \chi^e(f,k_e) + \tau_e\delta k_e - c\Big] \\ &+ V_a(b,a,k_e,y_l)(r^a a + d) + V_k(b,a,k_e,y_l)(f - \delta k_e) \\ &+ \frac{1}{2}k_e\sigma_k^2 V_{kk}(b,a,k_e,y_l) \\ &+ \lambda_{y,lh}(V(b,a,k_e,y_h) - V(b,a,k_e,y_l))\end{aligned} \tag{3.12}$$

and analogously for $V(b,a,k_e,y_h)$, i.e. high productivity types.

The solution to the workers' problem is characterized by

$$\begin{aligned}(\rho+\zeta)\,V(b,a,z) &= \max_{c,\ell,d} u(c,\ell) + V_b(b,a,z)\Big[(1-\tau)w\exp(z)\ell + r^b(b)b + T - d - \chi^a(d,a) - c\Big] \\ &+ V_a(b,a,z)(r^a a + d) \qquad\qquad (3.13) \\ &+ \sum_{j\in\{1,2\}} V_{z_j}(b,a,z)(-\beta_j z_j) + \lambda_j \int_{-\infty}^{\infty} (V(b,a,x) - V(b,a,z_j))\phi_j(x)\mathrm{d}x\end{aligned}$$

where $\phi_j(x)$ denotes the pdf of a normal distribution with standard deviation σ_{zj}.

C.3 Derivation of Phillips Curve

Here we derive the New Keynesian Phillips curve following Appendix B.2 of KMV. In recursive form, the profit maximization problem of an intermediate goods producer can be written as

$$\begin{aligned}r_t^a J(t,P_t(j)) = \max_{\pi_{jt}} \left(\frac{P_t(j)}{P_t}\right)^{1-\epsilon} Y_t - mc_t\left(\frac{P_t(j)}{P_t}\right)^{-\epsilon} Y_t - \frac{\theta}{2}\pi_{jt}^2 Y_t \\ + J_t(t,P_t(j)) + P_t(j)\pi_{jt}J_p(t,P_t(j)),\end{aligned} \tag{3.14}$$

where π_{jt} is the firm-specific inflation rate $\pi_{jt} = \frac{\dot{P}_t(j)}{P_t(j)}$. The first order condition for the maximization problem in (3.14) is

$$J_p(t,P_t(j)) = \frac{\theta\pi_{jt}Y_t}{P_t(j)}, \tag{3.15}$$

the envelope condition is

$$\begin{aligned}(r_t^a - \pi_{jt})J_p(t,P_t(j)) = \frac{Y_t}{P_t}(1-\epsilon)\left(\frac{P_t(j)}{P_t}\right)^{-\epsilon} + \frac{Y_t}{P_t}\epsilon\cdot mc_t\left(\frac{P_t(j)}{P_t}\right)^{-\epsilon-1} \\ + J_{tp}(t,P_t(j)) + \pi_{jt}P_t(j)J_{pp}(t,P_t(j))\;.\end{aligned}$$

Since firms are symmetric, $P_t(j) = P_t$ and $\pi_{jt} = \pi_t$, and we have

$$(r_t^a - \pi_t)J_p(t, P_t) = (1-\epsilon)\frac{Y_t}{P_t} + \epsilon \cdot mc_t \frac{Y_t}{P_t} + J_{tp}(t, P_t) + \pi_t P_t J_{pp}(t, P_t). \tag{3.16}$$

Taking the derivative of (3.15) with respect to time gives

$$J_{pp}(t, P_t(j))\dot{P}_t(j) + J_{tp}(t, P_t(j)) = \frac{\theta \dot{\pi}_t Y_t}{P_t(j)} + \frac{\theta \pi_t \dot{Y}_t}{P_t(j)} - \frac{\theta \pi_t Y_t}{P_t(j)} \frac{\dot{P}_t(j)}{P_t(j)}.$$

Recall that $\pi_t P_t = \dot{P}_t$ and substitute the above expression into (3.16). This gives the New Keynesian Phillips curve (3.4)

$$\left(r_t^a - \frac{\dot{Y}_t}{Y_t}\right)\pi_t = \frac{\epsilon}{\theta}\left[mc_t - \frac{\epsilon - 1}{\epsilon}\right] + \dot{\pi}_t.$$

D Monetary Policy Shocks and Portfolio Shares

D.1 Portfolio shares

This section details how we arrive at our estimates for the (log of the) portfolio shares at different percentiles p of the business return distribution. We closely follow Luetticke (2021) in arriving at these estimates.

We first sort entrepreneurial households in a given year t by their business return r^e. Next, we calculate the percentile of each household in the return distribution as

$$prctl_i = \frac{\sum_{j: r_j^e < r_i^e} w_j}{\sum_j w_j}$$

where w_j denote the sample weights provided by the SCF. For each percentile, we then regress the log of the portfolio share on the appropriately adjusted percentile measures. Specifically, to estimate the portfolio share at the p'th percentile, we perform a weighted regression

$$\ln(\text{portf. share}_i)\, \omega_i = \alpha\, \omega_i + \beta(prctl_i - p)\, \omega_i + u_i$$

where u is an error term and the weight we use for observation i is

$$\omega_i = \sqrt{w_i \phi\left(\frac{prctl_i - p}{0.1}\right)},$$

where $\phi(\cdot)$ corresponds to the probability density function of a standard normal distribution. The estimate of the intercept α is our estimate of the log of the portfolio share at percentile p for the year t.

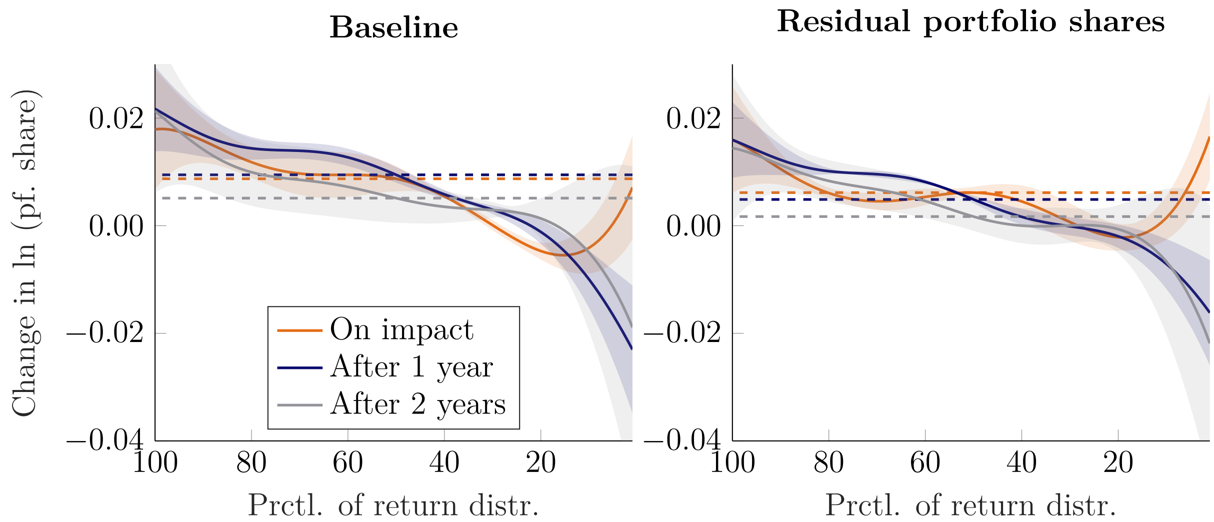


Figure 3.18: Impulse responses of portfolio shares to Gertler and Karadi (2015) monetary policy shock.
Notes: Change in the logarithm of portfolio shares following a 25 basis points expansionary monetary policy shock by business return percentile. The dashed lines depict the responses at the median of the return distribution. Confidence bands are at the 66% level.

D.2 Gertler and Karadi (2015) shocks

In our baseline specification we used the Romer & Romer shock series as instruments for the federal funds rate. Here, we instead employ the shock series derived by Gertler and Karadi (2015) who use high-frequency data to identify monetary policy surprises. We follow Ramey (2016) in focusing on the series that uses the 3-month ahead fed funds futures as instruments. Figure 3.18 shows the results.

The median responses are very similar to those found when using the Romer & Romer shocks series. We find the robustness of the results in this regard encouraging, given that the time window that is covered by the Gertler and Karadi (2015) shocks (and hence the SCF waves that we use to estimate portfolio shares) has only a small overlap with that of the Romer & Romer shocks. Regarding the heterogeneity of responses, the u-shape that we find when using the Romer & Romer shocks only appears in the impact response to the shock, and most clearly when using the residual portfolio shares.

E Additional Graphs and Tables

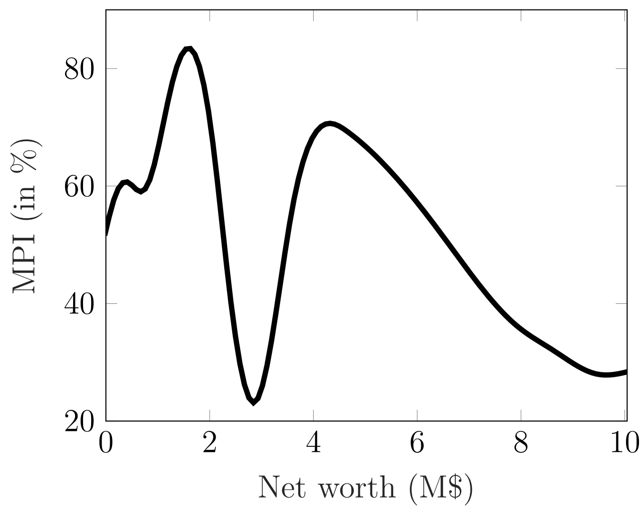


Figure 3.19: MPI into the private business out of a transfer of $ 500 into the liquid account over one quarter.

Table 3.17: Summary statistics SCF since 1989

	mean	p25	p50	p75	sd	min	max
Business return	0.60	0.05	0.19	0.60	1.04	-0.89	6.73
Net worth	30.0	0.9	4.3	20.2	90.4	0.0	1861.6
Bus. wealth	17.0	0.3	1.3	8.4	62.9	0.0	1300.6
Observations	7288						

Notes: Net worth and business wealth are measured in million US dollars, all deflated to 2019. Sample selection is as explained in main text. The displayed summary statistics do not make use of the sampling weights. Therefore, as the SCF oversamples rich individuals, the means and percentiles of net worth and business wealth appear large in comparison to equivalent statistics generated from the model (see for a comparison Figure 3.6). In all other analyses we conduct in this paper, we use the sampling weights provided by the SCF.

Bibliography

Bibliography

Acemoglu, Daron and David H Autor (2011). *Skills, tasks and technologies: Implications for employment and earnings*. Vol. 4. PART B. Elsevier B.V., pp. 1043–1171.

Acemoglu, Daron, Andrea Manera, and Pascual Restrepo (2020). *Does the US Tax Code Favor Automation?* NBER Working Papers 27052. National Bureau of Economic Research, Inc.

Acemoglu, Daron and Pascual Restrepo (2018a). "Modeling Automation". In: *AEA Papers and Proceedings* 108, pp. 48–53.

— (2018b). "The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment". In: *American Economic Review* 108.6, pp. 1488–1542.

Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll (2020). "Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach".

Adam, Klaus and Junyi Zhu (2016). "Price-Level Changes And The Redistribution Of Nominal Wealth Across The Euro Area". In: *Journal of the European Economic Association* 14.4, pp. 871–906.

Adão, Rodrigo, Martin Beraja, and Nitya Pandalai-Nayar (2020). *Technological Transitions with Skill Heterogeneity Across Generations*. NBER Working Papers 26625. National Bureau of Economic Research, Inc.

Ahn, Sehyoun, Greg Kaplan, Benjamin Moll, Thomas Winberry, and Christian Wolf (2018). "When Inequality Matters for Macro and Macro Matters for Inequality". In: *NBER Macroeconomics Annual 2017, Volume 32*. University of Chicago Press. Chap. 1, pp. 1–75.

Aiyagari, S. Rao (1994). "Uninsured Idiosyncratic Risk and Aggregate Saving". In: *Quarterly Journal of Economics* 109.3, pp. 659–684.

Albanesi, Stefania and Jiyeon Kim (2021). *The Gendered Impact of the COVID-19 Recession on the US Labor Market*. NBER Working Papers 28505. National Bureau of Economic Research, Inc.

Almgren, Mattias, Jose E Gallegos, John Kramer, and Ricardo Lima (2020). "Monetary Policy and Liquidity Constraints: Evidence from the Euro Area".

Alpanda, Sami and Sarah Zubairy (2019). "Household Debt Overhang and Transmission of Monetary Policy". In: *Journal of Money, Credit and Banking* 51.5, pp. 1265–1307.

Altavilla, Carlo, Luca Brugnolini, Refet S. Gürkaynak, Roberto Motto, and Giuseppe Ragusa (2019). “Measuring euro area monetary policy”. In: *Journal of Monetary Economics* 108, pp. 162–179.

Alvaredo, Facundo, Anthony B Atkinson, Salvatore Morelli, and Salvatore Morelli (2016). *Top wealth shares in the UK over more than a century.* WID.world Working Paper Series 2017/2.

Alvaredo, Facundo, Lucas Chancel, Thomas Piketty, Emmanuel Saez, and Gabriel Zucman (2018). *The World Inequality Report.* Harvard University Press.

Alves, Felipe, Greg Kaplan, Benjamin Moll, and Giovanni L. Violante (2020). “A Further Look at the Propagation of Monetary Policy Shocks in HANK”. In: *Journal of Money, Credit and Banking* 52.S2, pp. 521–559.

Andersen, Asger Lau, Niels Johannesen, Mia Jørgensen, and José-Luis Peydró (2021). *Monetary Policy and Inequality.* EconStor Preprints 227763. ZBW - Leibniz Information Centre for Economics.

Ascari, Guido and Timo Haber (2021). “Non-linearities, state-dependent prices and thetransmission mechanism of monetary policy”.

Asker, John, Allan Collard-Wexler, and Jan De Loecker (2014). “Dynamic Inputs and Resource (Mis)Allocation”. In: *Journal of Political Economy* 122.5, pp. 1013–1063.

Asker, John, Joan Farre-Mensa, and Alexander Ljungqvist (2015). “Corporate Investment and Stock Market Listing: A Puzzle?” In: *The Review of Financial Studies* 28.2, pp. 342–390.

Auclert, Adrien (2019). “Monetary Policy and the Redistribution Channel”. In: *American Economic Review* 109.6, pp. 2333–2367.

Auclert, Adrien, Matthew Rognlie, and Ludwig Straub (2018). *The Intertemporal Keynesian Cross.* NBER Working Papers 25020. National Bureau of Economic Research, Inc.

— (2020). “Micro Jumps, Macro Humps: Monetary Policy and Business Cycles in an Estimated HANK Model”.

Auerbach, Alan J and Yuriy Gorodnichenko (2013). “Fiscal Multipliers in Recession and Expansion”. In: *Fiscal Policy after the Financial Crisis.* Ed. by Alberto Alesina and Francesco Giavazzi. Chicago, IL: University of Chicago Press, pp. 63–98.

Autor, David H and David Dorn (2013). “The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market”. In: *American Economic Review* 103.5, pp. 1553–1597.

Autor, David H, Frank Levy, and Richard J Murnane (2003). “The Skill Content of Recent Technological Change: An Empirical Exploration”. In: *The Quarterly Journal of Economics* November, pp. 1279–1333.

Bachmann, Ronald, Merve Cim, and Colin Green (2019). “LongRun Patterns of Labour Market Polarization: Evidence from German Micro Data”. In: *British Journal of Industrial Relations* 57.2, pp. 350–376.

Bahaj, Saleem, Angus Foulis, Gabor Pinter, and Paolo Surico (2020). "Employment and the Collateral Channel of Monetary Policy".

Barnes, Aaron (2021). *Personal Wealth, 2016.* Tech. rep. Spring. Internal Revenue Service.

Benhabib, Jess, Alberto Bisin, and Shenghao Zhu (2015). "The wealth distribution in Bewley economies with capital income risk". In: *Journal of Economic Theory* 159, pp. 489–515.

Beraja, Martin, Andreas Fuster, Erik Hurst, and Joseph Vavra (2019). "Regional Heterogeneity and the Refinancing Channel of Monetary Policy". In: *The Quarterly Journal of Economics* 134.1, pp. 109–183.

Bessen, James, Maarten Goos, Anna Salomons, and Wiljan van den Berge (2020). "What Happens to Workers at Firms that Automate?"

Bewley, Truman (1983). "A Difficulty with the Optimum Quantity of Money". In: *Econometrica* 51.5, pp. 1485–1504.

Bilbiie, Florin O. (2008). "Limited asset markets participation, monetary policy and (inverted) aggregate demand logic". In: *Journal of Economic Theory* 140.1, pp. 162–196.

— (2020). "The New Keynesian cross". In: *Journal of Monetary Economics* 114.C, pp. 90–108.

Bilbiie, Florin O., Diego R. Känzig, and Paolo Surico (2020). "Capital, Income Inequality, and Consumption: the Missing Link".

BIS (2021). "BIS Annual Economic Report 2021". In: Bank of International Settlements. Chap. The distributional footprint of monetary policy, pp. 39–64.

Brinca, Pedro, Hans A. Holter, Per Krusell, and Laurence Malafry (2016). "Fiscal multipliers in the 21st century". In: *Journal of Monetary Economics* 77, pp. 53–69.

Cagetti, Marco and Mariacristina De Nardi (2006). "Entrepreneurship, Frictions, and Wealth". In: *Journal of Political Economy* 114.5, pp. 835–870.

Carlino, Gerald and Robert DeFina (1998). "The Differential Regional Effects of Monetary Policy". In: *Review of Economics and Statistics* 80.4, pp. 572–587.

— (1999). "The differential regional effects of monetary policy: Evidence from the US states". In: *Journal of Regional science* 39.2, pp. 339–358.

Chernoff, Alex W. and Casey Warman (2020). *COVID-19 and Implications for Automation.* NBER Working Papers 27249. National Bureau of Economic Research, Inc.

Chetty, Raj, John N Friedman, Nathaniel Hendren, Michael Stepner, and The Opportunity Insights Team (2020). *The Economic Impacts of COVID-19: Evidence from a New Public Database Built Using Private Sector Data.* Working Paper 27431. National Bureau of Economic Research.

Chow, Gregory C and An-loh Lin (1971). "Best Linear Unbiased Interpolation, Distribution, and Extrapolation of Time Series by Related Series". In: *The Review of Economics and Statistics* 53.4, pp. 372–375.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans (1999). "Monetary policy shocks: What have we learned and to what end?" In: *Handbook of Macroeconomics*. Ed.

by John B Taylor and Michael Woodford. Vol. 1. PART A. Amsterdam: Elsevier. Chap. 2, pp. 65–148.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans (2005). “Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy”. In: *Journal of Political Economy* 113.1, pp. 1–45.

Cloyne, James, Clodomiro Ferreira, Maren Froemel, and Paolo Surico (2020). “Monetary Policy, Corporate Finance and Investment”.

Cloyne, James and Patrick Hürtgen (2016). “The macroeconomic effects of monetary policy: A new measure for the United Kingdom”. In: *American Economic Journal: Macroeconomics* 8.4, pp. 75–102.

Coibion, Olivier (2012). “Are the effects of monetary policy shocks big or small?” In: *American Economic Journal: Macroeconomics* 4.2, pp. 1–32.

Coibion, Olivier, Yuriy Gorodnichenko, Lorenz Kueng, and John Silvia (2017). “Innocent bystanders? Monetary policy and inequality”. In: *Journal of Monetary Economics* 88, pp. 70–89.

Cooley, Thomas F. and Vincenzo Quadrini (2001). “Financial Markets and Firm Dynamics”. In: *American Economic Review* 91.5, pp. 1286–1310.

Cortes, Guido Matias (2016). “Where Have the Middle-Wage Workers Gone? A Study of Polarization Using Panel Data”. In: *Journal of Labor Economics* 34.1.

Cortes, Guido Matias and Giovanni Gallipoli (2018). “The Costs of Occupational Mobility: An Aggregate Analysis”. In: *Journal of the European Economic Association* 16.2, pp. 275–315.

Cortes, Guido Matias, Nir Jaimovich, Christopher J. Nekarda, and Henry E. Siu (2020). “The dynamics of disappearing routine jobs: A flows approach”. In: *Labour Economics* 65, p. 101823.

Cortes, Guido Matias, Nir Jaimovich, and Henry E. Siu (2017). “Disappearing routine jobs: Who, how, and why?” In: *Journal of Monetary Economics* 91, pp. 69–87.

Dauth, Wolfgang, Sebastian Findeisen, Jens Suedekum, and Nicole Woessner (2021). “The Adjustment of Labor Markets to Robots”. In: *Journal of the European Economic Association.* jvab012.

De Nardi, Mariacristina, Phil Doctor, and Spencer D Krane (2007). “Evidence on entrepreneurs in the United States: Data from the 1989 2004 Survey of Consumer Finances”. In: *Economic Perspectives* 31.4, pp. 18–36.

DeBacker, Jason, Vasia Panousi, and Shanthi Ramnath (2018). “A Risky Venture: Income Dynamics within the Non-CorporatePrivate Business Sector”.

Dempsey, Kyle and Felicia Ionescu (2021). *Lending Standards and Borrowing Premia in Unsecured Credit Markets.* Tech. rep.

Dorn, David (2009). “Essays on Inequality, Spatial Interaction, and the Demand for Skills”. Doctoral dissertation. University of St. Gallen.

Dossche, Maarten, Jiri Slacalek, and Guido Wolswijk (2021). “Monetary policy and inequality”. In: *Economic Bulletin Articles* 2.

Dvorkin, Maximiliano and Alexander Monge-Naranjo (2019). *Occupation Mobility, Human Capital and the Aggregate Consequences of Task-Biased Innovations*. Tech. rep.

Eden, Maya and Paul Gaggl (2018). “On the Welfare Implications of Automation”. In: *Review of Economic Dynamics* 29, pp. 15–43.

Edin, Per-Anders, Tiernan Evans, Georg Graetz, Sofia Hernnäs, and Guy Michaels (2021). “Individual Consequences of Occupational Decline”.

Fairlie, Robert (2020). “The impact of COVID19 on small business owners: Evidence from the first three months after widespread socialdistancing restrictions”. In: *Journal of Economics & Management Strategy* 29.4, pp. 727–740.

Gathmann, Christina and Uta Schönberg (2010). “How General Is Human Capital? A Task-Based Approach”. In: *Journal of Labor Economics* 28.1.

Gertler, Mark and Peter Karadi (2015). “Monetary Policy Surprises, Credit Costs,and Economic Activity”. In: *American Economic Journal: Macroeconomics* 7.1, pp. 44–76.

Goos, Maarten, Alan Manning, and Anna Salomons (2014). “Explaining Job Polarization: Routine-Biased Technological Change and Offshoring”. In: *American Economic Review* 104.8, pp. 2509–2526.

Gornemann, Nils, Keith Kuester, and Makoto Nakajima (2016). “Doves for the Rich, Hawks for the Poor? Distributional Consequences of Monetary Policy”. In: *International Finance Discussion Paper* 2016.1167, pp. 1–40.

Granja, João, Christos Makridis, Constantine Yannelis, and Eric Zwick (2020). *Did the Paycheck Protection Program Hit the Target?* Working Paper 27095. National Bureau of Economic Research.

Greenwood, Jeremy, Zvi Hercowitz, and Per Krusell (1997). “Long-Run Implications of Investment-Specific Technological Change”. In: *American Economic Review* 87.3, pp. 342–362.

Gutiérrez, Germán and Thomas Philippon (2017). “Investmentless Growth: An Empirical Investigation”. In: *Brookings Papers on Economic Activity* 48.2 (Fall), pp. 89–190.

Guvenen, Fatih, Fatih Karahan, Serdar Ozkan, and Jae Song (2021). “What Do Data on Millions of U.S. Workers Reveal About Lifecycle Earnings Dynamics?” In: *Econometrica* 89.5, pp. 2303–2339.

Hershbein, Brad and Lisa B. Kahn (2018). “Do Recessions Accelerate Routine-Biased Technological Change? Evidence from Vacancy Postings”. In: *American Economic Review* 108.7, pp. 1737–1772.

Hubbard, R. Glenn and Michael R Strain (2020). *Has the Paycheck Protection Program Succeeded?* Working Paper 28032. National Bureau of Economic Research.

Hubmer, Joachim, Per Krusell, and Anthony A. Smith (2021). “Sources of US Wealth Inequality: Past, Present, and Future”. In: *NBER Macroeconomics Annual* 35.1, pp. 391–455.

Huggett, Mark (1993). "The risk-free rate in heterogeneous-agent incomplete-insurance economies". In: *Journal of Economic Dynamics and Control* 17.5-6, pp. 953–969.

Jacobson, Darien B., Brian G. Raub, and Barry W. Johnson (2007). *The Estate Tax: Ninety Years and Counting*. Tech. rep. Summer. Internal Revenue Service.

Jaimovich, Nir, Itay Saporta-Eksten, Henry E. Siu, and Yaniv Yedid-Levi (2021). "The Macroeconomics of Automation: Data, Theory, and Policy Analysis".

Jaimovich, Nir and Henry E. Siu (2020). "Job Polarization and Jobless Recoveries". In: *The Review of Economics and Statistics* 102.1, pp. 129–147.

Jeenas, Priit (2019). "Firm Balance Sheet Liquidity, Monetary Policy Shocks, and Investment Dynamics".

Jordà, Òscar (2005). "Estimation and Inference of Impulse Responses by Local Projections". In: *American Economic Review* 95.1, pp. 161–182.

Jung, Jaewon and Jean Mercenier (2014). "Routinization-Biased Technical Change And Globalization: Understanding Labor Market Polarization". In: *Economic Inquiry* 52.4, pp. 1446–1465.

Kambourov, Gueorgui and Iourii Manovskii (2009a). "Occupational Mobility and Wage Inequality". In: *Review of Economic Studies* 76.2, pp. 731–759.

— (2009b). "Occupational specificity of human capital". In: *International Economic Review* 50.1, pp. 63–115.

Kaplan, Greg, Benjamin Moll, and Giovanni L. Violante (2018). "Monetary Policy According to HANK". In: *American Economic Review* 108.3, pp. 697–743.

— (2020). *The Great Lockdown and the Big Stimulus: Tracing the Pandemic Possibility Frontier for the U.S.* NBER Working Papers 27794. National Bureau of Economic Research, Inc.

Kaplan, Greg, Giovanni L. Violante, and Justin Weidner (2014). "The Wealthy Hand-to-Mouth". In: *Brookings papers on Economic Activity* Spring, pp. 77–153.

Kartashova, Katya and Xiaoqing Zhou (2021). "Wealth Inequality and Return Heterogeneity during the COVID-19 Pandemic".

Kikuchi, Shinnosuke and Sagiri Kitao (2020). *Welfare Effects of Polarization: Occupational Mobility over the Life-cycle*. Tech. rep.

Kopczuk, Wojciech (2015). "What Do We Know about the Evolution of Top Wealth Shares in the United States?" In: *Journal of Economic Perspectives* 29.1, pp. 47–66.

Kopczuk, Wojciech and Emmanuel Saez (2004). "Top Wealth Shares in the United States, 1916-2000: Evidence from Estate Tax Returns". In: *National Tax Journal* 57.2, Part 2, pp. 445–487.

Krusell, Per and Anthony Smith (1998). "Income and Wealth Heterogeneity in the Macroeconomy". In: *Journal of Political Economy* 106.5, pp. 867–896.

Kuhn, Moritz, Moritz Schularick, and Ulrike I. Steins (2020). "Income and Wealth Inequality in America, 19492016". In: *Journal of Political Economy* 128.9, pp. 3469–3519.

Leahy, John V. and Aditi Thapar (2019). “Demographic Effects on the Impact of Monetary Policy”.

Levine, Ross and Yona Rubinstein (2020). *Selection into entrepreneurship and self-employment.* CEPR Discussion Papers 15143. C.E.P.R. Discussion Papers.

Lochner, Lance and Alexander Monge-Naranjo (2012). “Credit Constraints in Education”. In: *Annual Review of Economics* 4.1, pp. 225–256.

— (2016). “Student Loans and Repayment: Theory, Evidence, and Policy”. In: vol. 5. Handbook of the Economics of Education. Elsevier. Chap. 8, pp. 397–478.

Lucas, Robert (1978). “On the Size Distribution of Business Firms”. In: *Bell Journal of Economics* 9.2, pp. 508–523.

Luetticke, Ralph (2021). “Transmission of Monetary Policy with Heterogeneity in Household Portfolios”. In: *American Economic Journal: Macroeconomics* 13.2, pp. 1–25.

Meier, Matthias and Timo Reinelt (2020). *Monetary Policy, Markup Dispersion, and Aggregate TFP*. CRC TR 224 Discussion Paper Series. University of Bonn and University of Mannheim, Germany.

Melcangi, Davide and Vincent Sterk (2020). *Stock Market Participation, Inequality, and Monetary Policy.* Staff Reports 932. Federal Reserve Bank of New York.

Moll, Benjamin, Lukasz Rachel, and Pascual Restrepo (2021). *Uneven Growth: Automation’s Impact on Income and Wealth Inequality.* Working Paper 28440. National Bureau of Economic Research.

Mukoyama, Toshi, Satoshi Tanaka, and Naoki Takayama (2021). “Occupational Reallocations within and across Firms: Implications for the Labor Market Polarization”.

Newey, Whitney K and Kenneth D West (1987). “A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix”. In: *Econometrica* 55.3, pp. 703–708.

OECD (2020). *OECD Indicators of Employment Protection.* OECD. URL: `www.oecd.org/employment/protection`.

Ottonello, Pablo and Thomas Winberry (2020). “Financial Heterogeneity and the Investment Channel of Monetary Policy”. In: *Econometrica* 88.6, pp. 2473–2502.

Owyang, Michael T. and Howard J. Wall (2009). “Regional VARs and the channels of monetary policy”. In: *Applied Economics Letters* 16.12, pp. 1191–1194.

Patterson, Christina (2019). “The Matching Multiplier and the Amplification of Recessions”.

Piketty, Thomas (2014). *Capital in the Twenty-First Century:* Harvard University Press.

Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman (2018). “Distributional National Accounts: Methods and Estimates for the United States”. In: *The Quarterly Journal of Economics* 133.2, pp. 553–609.

Poschke, Markus (2013). “Who becomes an entrepreneur? Labor market prospects and occupational choice”. In: *Journal of Economic Dynamics and Control* 37.3, pp. 693–710.

Ramey, Valerie A. (2016). “Macroeconomic shocks and their propagation”. In: *Handbook of Macroeconomics*. Ed. by Harald Uhlig and John B Taylor. 1st ed. Vol. 2. Amsterdam: Elsevier B.V. Chap. 3, pp. 71–162.

Ramey, Valerie A. and Sarah Zubairy (2018). “Government Spending Multipliers in Good Times and in Bad: Evidence from US Historical Data”. In: *Journal of Political Economy* 126.2, pp. 850–901.

Romer, Christina and David Romer (2004). “A new measure of monetary shocks: derivation and implications”. In: *American Economic Review* 94.4, pp. 1055–1084.

Rotemberg, Julio J. (1982). “Monopolistic Price Adjustment and Aggregate Output”. In: *The Review of Economic Studies* 49.4, p. 517.

Saez, Emmanuel and Gabriel Zucman (2016). “Wealth inequality in the United States since 1913: Evidence From Capitalized Income Tax Data”. In: *Quarterly Journal of Economics* 131.2, pp. 519–578.

Schwartz, Marvin (1988). *Estimates of Personal Wealth, 1982—A Second Look*. Tech. rep. Spring. Internal Revenue Service.

Slacalek, Jiri, Oreste Tristani, and Giovanni L. Violante (2020). “Household balance sheet channels of monetary policy: A back of the envelope calculation for the euro area”. In: *Journal of Economic Dynamics and Control* 115.C.

Stock, James H. and Mark W. Watson (2018). “Identification and Estimation of Dynamic Causal Effects in Macroeconomics Using External Instruments”. In: *Economic Journal* 128.610, pp. 917–948.

Sullivan, Paul (2010). “Empirical evidence on occupation and industry specific human capital”. In: *Labour Economics* 17.3, pp. 567–580.

Tan, Eugene (2020). “Entrepreneurial Investment Dynamics and the Wealth Distribution”.

Tenreyro, Silvana and Gregory Thwaites (2016). “Pushing on a String: US Monetary Policy Is Less Powerful in Recessions”. In: *American Economic Journal: Macroeconomics* 8.4, pp. 43–74.

vom Lehn, Christian (2020). “Labor Market Polarization, the Decline of Routine Work, and Technological Change: A Quantitative Analysis”. In: *Journal of Monetary Economics* 110, pp. 62–80.

Werning, Ivan (2015). *Incomplete Markets and Aggregate Demand*. Tech. rep. October, pp. 1–46.

Wong, Arlene (2021). “Refinancing and the Transmission of Monetary Policy to Consumption”.

Xavier, Inês (2021). *Wealth Inequality in the US: the Role of Heterogeneous Returns*.