Essays in Macroeconomics and Agents' Heterogeneity

Marta Morazzoni

TESI DOCTORAL UPF / 2023

THESIS SUPERVISORS

Dr. Isaac Baley
Dr. Edouard Schaal

Department of Economics and Business

To my best friends

My friends are my estate.

Forgive me then the avarice to hoard them.

They tell me those who were poor early have different views of gold.

I don't know how that is.

God is not so wary as we, else He would give us no friends,

lest we forget Him.

Emily Dickinson, 1858

Acknowledgments

I thank my irreplaceable and most brilliant advisors, Edouard Schaal and Isaac Baley, for having accompanied and guided me throughout my journey as a graduate student. Thank you for your generous patience, your invaluable feedback and your constant encouragement. Thank you for teaching me how to be a thoughtful economist, for the kind and understanding smile with which you have reasoned me out of countless PhD-life crises, and for showing me, with your direct example, how to be a constructive and supporting researcher. I will never cease to be your student.

I carry in my heart several professors who were never official advisors but almost felt like ones. In particular, I thank Andrea Caggese, Jan Eeckhout, Libertad Gonzalez and Priit Jeenas for always keeping their door open and for uncountable suggestions over the years. And I have the most sincere gratitude for Jaume Ventura, because after the first 10 minutes of the first lecture he gave us six years ago, I knew I wanted to be an economist: thank you for always pushing me to ask the difficult questions. Finally, my PhD journey would have not been the same without Mariacristina De Nardi. Thank you for believing I had something to say, for making me think deeply about the implications of my work, and for being a great role model, especially for all the women in Macro.

I would not be half the person and the researcher I am without the most amazing colleagues I could have wished for. To the UPF crew, and especially to Ayah Bohsali, Josefina Cenzón, Shubhdeep Deb, Pia Ennusschat, Malachy Gavan, Cristiano Mantovani, Claudia Meza-Quadra, Maria Ptashkina and Alejandro Rábano. To Guido Bongioanni and Lukas Nord at the European University Institute, and to my loving Minnesotan family – Lidia Cruces, Nicolo Russo, Ante Sterc, Jacob Strauss and Katherina Thomas. I feel blessed to have shared these years with you.

One of the greatest gifts this PhD made me is the friendship with Laia Bosque, Marta Cota, Valeria Gargiulo, Vladislav Morozov, Laura Perez, Danila Smirnov and Andrea Sy, who have been and are my home, my anchor and my light. You deserve all the wonders this universe has to offer (and more!), and you will change this world for the better. Thank you for always walking by my side.

My unconditional love goes to my family, which is my blood and my fierce identity. To my mum, for her delicate strength, like the one of the tulips she so much adores. To my dad, an incurable optimist, for all the lullabies he invented when I had nightmares. And to my grandma and aunts, for the tenacious resilience that has never abandoned them through their life. *A Mamma Robi, Papà Paolo, Nonna Igia, Zia Anna e Zia Bombo, che ci sono e ci saranno fino alla fine del mondo.*

Words fell short to express my gratitude to my partner Marco. Thank you for your healthy love, for the carbonara(s) and the laughs you always had for me during my job market, for caring and listening to my dreams and fears, for helping me blossom again, for bearing my bad taste for movies and my sloppiness with everyday practicalities, and for being you, a smart, generous and genuine person. For all the times we sang our song driving under the dark starry skies of Lanzarote, and for the courage and the love with which we choose to be in each other's lives every day.

My thesis is dedicated to my friends, who are the lymph of my existence. To Carla and Andrea, who are pure joy. To Lube, Gio and Isa, because, even if far apart, we always find a way to grab each other's hand: thank you for all the adventures and for your unwavering affection. To Erika, Meli and Paggi, who are and would be my best friends in this life and in a thousand more: thank you for being my enduring roots. And to Giuli, Cami, Franci, Chia and Sibi, because every time I see a group of old grannies I think about us, about all the steps we will give together, and I feel proud of what each of us has accomplished. My dear friends, you are and will forever be my heart.

Abstract

In my PhD thesis, I analyse different macro and policy-relevant heterogeneities across both firms and individuals. In the midst of vivid debate over the US college financial aid system, Chapter 1 shows that student loans can affect entrepreneurial entry, capital allocation and aggregate productivity in the US. Chapter 2 uncovers instead a novel empirical fact, namely that dominant firms in the US have a lower passthrough from input to output prices after an interest rate shock, and rationalizes it in a New Keynesian model where heterogeneous markup responses are due to the different demand elasticities faced by firms over their life cycle. In Chapter 3, I endogenize family formation in a joint-search model, and argue that understanding how and when agents match in the marriage market is key to replicate the empirically-estimated differences in wages and unemployment rates by marital status, and it is a key phenomenon to understand in order to design an optimal unemployment public insurance.

Resumen

En mi tesis doctoral, examino diversas heterogeneidades en empresas e individuos que son relevantes para las políticas públicas y las variables macroeconómicas. En medio del vívido debate sobre el sistema de ayuda financiera para universidades de EE.UU., el Capítulo 1 muestra que los préstamos estudiantiles pueden afectar la formación de empresas, la asignación de capital y la productividad agregada en EE.UU. El Capítulo 2 descubre que las empresas dominantes en los EE.UU. tienen un traspaso más bajo de los costes de los insumos a los precios de los productos después de un shock en la tasa de interés, y lo racionaliza en un modelo neokeynesiano donde dichas respuestas heterogéneas se deben a las diferentes elasticidades de la demanda que enfrentan las empresas a lo largo de su ciclo de vida. En el Capítulo 3, endogenizo la formación de familias en un modelo de búsqueda conjunta en el mercado laboral, para explicar las diferencias empíricas en salarios y tasas de desempleo por estado civil y diseñar un seguro público de desempleo óptimo.

Preface

The topic of my thesis is the analysis of macro and policy-relevant heterogeneities, across both firms and individuals. In particular, by studying barriers to agents' mobility across the income distribution, my research focuses on how demographic factors and heterogeneous life-cycle experiences shape occupation choices, and how this in turn affects aggregate outcomes such as business dynamism, the efficiency of resource allocation and total production. To investigate the macro-level implications of micro-level heterogeneities, I contribute to the existing literature by developing quantitative models that are supported by the econometric analysis of firm and household-level data, and that can serve as laboratories for policy counterfactuals to inform the public debate.

In the first chapter of my thesis, I study the interplay of education choices and entrepreneurial outcomes over the life-cycle of individuals, focusing on the effect of student debt on business entry and funding. As of today, 1 in 4 US labor force participants has borrowed to finance their degree, and student loans have become the second largest debt market in the country, stirring a vivid debate on the soaring costs of higher education. While college borrowing is a key form of financial aid, there is little understanding as to whether and how it may interact with the entrepreneurial margins of indebted college graduates. Since higher education is generally associated with better business outcomes, student loans could then be consequential for overall firm dynamism and aggregate quantities. I tackle these questions by gathering novel evidence from a large US dataset, which I use to motivate and discipline a macroeconomic model that can serve as a quantitative framework to evaluate highly-debated government interventions. Specifically, I estimate the aggregate and distributional consequences of policy reforms such as the expansion of grants schemes or college borrowing limits, and the adoption of income-based repayment plans for student loans.

In the second chapter of my thesis, I instead combine elements from the literature on Family Economics in a macroeconomic analysis of labor market outcomes. In particular, joint with my colleague Danila Smirnov, I quantitatively study the interplay of job search decisions and family dynamics. Developing a novel heterogeneous agents model that combines joint-search and endogenous household's formation, we highlight how marital sorting and selection into joint-households affect wages and unemployment rates. Both mechanisms determine productivity differences in the sample composition of married and singles, which are crucial to replicate salient US labor market heterogeneities by marital status that were previously not accounted for. Our calibrated framework explains 75% of the wage marital premium, 60% of the unemployment marital gap and the bulk of the marital patterns documented in the US, and it is used as a laboratory to evaluate optimal unemployment insurance schemes and family benefits for single and married households.

In the third chapter of my thesis, together with Andrea Chiavari and Danila Smirnov, we explore the role of market power for heterogeneity in monetary policy transmission. Leveraging a US firm-level dataset including several industries, we measure markups using structural estimators from the empirical Industrial Organization literature, and document that dominant firms have a more countercyclical markup response after a contractionary monetary policy shock. Building a heterogeneous firms New Keynesian model with demand accumulation and endogenous markups, we show that this can be due to the different demand elasticities faced by firms. Dominant firms face a more inelastic demand, which implies a lower pass-through from input costs to output prices. Therefore, after a contractionary monetary policy shock, dominant firms pass less the reduction in marginal costs to prices compared to competitors, and increase their markups by more, as we document empirically. After calibrating our model to US data, we find that considering firms' heterogeneous demand elasticities has important implications for monetary policy amplification.

Contents

Part I

Student Debt and Entrepreneurship in the US

Marta Morazzoni*

1 Introduction

As of today, 60% of all college graduates and nearly 1 in 4 labor force participants in the US have borrowed to finance their degrees. Student loans have become the second largest debt market in the country – valued at 1.6 trillion dollars and worth 6% of the GDP – and are currently at the center of a vivid public debate. In particular, while college borrowing represents the main pathway to university for many US students, the cost of higher education has been rising faster than inflation and faster than the college premium over the past years. Such a steady increase in university prices has been accompanied by an unprecedented surge in the level of student debt per person. With the median borrower piling up more than 35K dollars of education loans, recent studies have documented far reaching implications of rising student debt for individuals' life choices, including their job search strategies, marital decisions and home-ownership rates.[1]

An aspect that has received less attention by the literature is whether and how student loans interact with entrepreneurial dynamics. In several countries – including the US – entrepreneurs play a pivotal role in enhancing job creation and innovation, and analyses of micro-level data reveal that firms started by college graduates tend to be bigger and exhibit higher returns and life-cycle growth (see Michelacci and Schivardi (2020) and Queiró (2021)).[2] While student loans can open the gate to higher education in the US, evidence from studies of household finance also suggests that they can lead young adults to suffer from debt overhang (see Di Maggio et al. (2019)). Since owners typically need funds to run their businesses, a natural question that follows is whether college borrowing could interact with the entrepreneurial decisions of indebted graduates. And if so, could this also have repercussions for US firm dynamism and macroeconomic aggregates?[3]

*Department of Economics and Business, Universitat Pompeu Fabra. Email: marta.morazzoni@upf.edu. I am grateful to my advisors Isaac Baley and Edouard Schaal for their constant guidance and support, and to Andrea Caggese, Mariacristina De Nardi, Priit Jeenas and Jaume Ventura for countless discussions that have improved substantially this paper. I thank Fil Babalievsky, Marta Cota, Lidia Cruces, Davide Debortoli, Jan Eeckhout, Jordi Gali, Albrecht Glitz, Libertad Gonzalez, Fatih Guvenen, William Jungerman, Loukas Karabarbounis, Gianmarco Leon, Alberto Martin, Laia Mercader, Jo Mullins, David Nagy, Lukas Nord, Laura Perez, Giacomo Ponzetto, Milan Quentel, Nicolo Russo, Danila Smirnov, Andrea Sy, Theodor Vladasel and participants to seminars at CREI and U of Minnesota for helpful comments. I acknowledge financial support from the Spanish Ministry of Economy and Competitiveness through the Severo Ochoa Programme for Centres of Excellence in R&D (CEX2019-000915-S). All errors are my own.

[1] See for example Alon et al. (2021), Luo and Mongey (2019), Folch and Mazzone (2020) and Abbott et al. (2019).

[2] Queiró (2021) uses the universe of firms and workers from Portugal to document that education affects firm size and growth, and that educated entrepreneurs seem to be specifically better at innovation and technology adoption. With data from the US, Michelacci and Schivardi (2020) further show that a high degree of complementarity between education and experience determines higher returns to entrepreneurship for graduate and post-graduate individuals.

[3] A recent survey by the US Chamber of Commerce found that 30% of small business owners think student loans have impacted their ability to grow their business. Moreover, the concern that student debt may constitute a barrier to entrepreneurship has long been a topic of political debate. For example, on June 30$^{\text{th}}$, 2021, Republican Nydia Velazquez, who is the chair of the House Small Business Committee, introduced a bill to establish a student loan debt forgiveness and deferment program for entrepreneurs (i.e: The 2021 Supporting America's Young Entrepreneurs Act).

This paper precisely intends to study the effects of student loans on occupational choices. The US student loans program allows individuals to afford higher education and potentially achieve higher productivity and earnings. But the key to college might not come without a cost, and student loans could in fact distort the entrepreneurial outcomes of indebted college graduates. To explore this perspective, I first provide novel empirical evidence on the link between education loans and entrepreneurial outcomes. I then build a quantitative model with interactions between college borrowing and entrepreneurial choices over the life-cycle of individuals, and analyse the importance of the US student loans program and its implications for firm creation, capital allocation and output. Finally, I use my framework as a laboratory to analyse potential reforms to university financing, such as expanding need- and merit-based grants, raising college borrowing limits, adopting income-driven repayment plans and cancelling outstanding education loans.

The first contribution of my work is to empirically document a negative relationship between college loans and the *extensive* and *intensive* margins of entrepreneurship. To this end, I leverage micro-level data from the Fed Survey of Consumer Finances (SCF) and focus on the 1989-2019 period. In the cross-section, higher levels of education are on average associated with better business outcomes for entrepreneurs, as shown in Michelacci and Schivardi (2020). Yet, individuals who took out student loans or carry higher outstanding balances at the time of the survey are less likely to become business owners and obtain funding compared to agents without a degree and to college graduates without student debt. Their firms are also relatively younger, tend to employ fewer workers and generate lower revenues and profits in absolute terms. Importantly, robustness analyses point towards the fact that my results are not systematically driven by negative selection into having student debt, both in terms of financial and individual productivity characteristics.

Several mechanisms could explain these findings: for the most part, education loans are settled through fixed repayment plans, carry a relatively high interest rate and cannot be discharged in bankruptcy, which may discourage or delay agents' entrepreneurial careers. Since outstanding liabilities can increase the chances of debt overhang, college borrowing could also tighten entrepreneurial financial constraints. Both factors may represent barriers to firm ownership, and, consistent with this observation, I show that indebted college graduates run businesses with higher profitability (relative to their size), which suggests they undergo a stricter selection at entry.

To rationalize my results, I develop a heterogeneous agents life-cycle framework, where individuals differ by wealth, productivity, age, education, and student debt. I build on the entrepreneurial models in Cagetti and De Nardi (2006) and Buera et al. (2011), but include an endogenous choice of college and student debt to analyse the interplay of education and entrepreneurial outcomes. Specifically, during youth, individuals decide whether to attend college or enter the labor market directly. University entails a tuition – net of grants – and students choose how much to take out in college debt, which is repaid after graduation. A unique feature of student loans is that they are *unsecured* credit granted by the Federal Government to a broad set of individuals, without requiring collateral and with the goal of reducing barriers to higher education. However, they are neither dischargeable nor collateralized by any physical asset one can borrow upon, which motivates the choice of modeling college borrowing and personal wealth separately.

In the model, education gives agents an income premium through a higher deterministic efficiency profile over their life-cycle, which is an incentive to enroll in college. Then, during adulthood, individuals make occupational choices and decide whether to open a firm or become workers. They save out of their income and consume a final good, which is produced by entrepreneurs combining their idiosyncratic productivity, capital and labor. In retirement, agents consume their pension and savings, and leave bequests. Finally, there is a government that collects income taxes, holds student debt and distributes grants and pensions. Equilibrium outcomes include the wage, the interest rate and the tax rate, and the general equilibrium (GE) setting of my framework allows me to study counterfactuals and policy reforms accounting for the full response of the economy.

The key contribution of the model is to link the dynamics of student debt and entrepreneurial choices through two main channels. First, the repayment of loans after graduation reduces the amount of resources individuals can save, and slows down wealth accumulation. Since personal assets are the collateral against which entrepreneurs borrow to finance capital acquisition, this mechanism has a negative effect on the entrepreneurial outcomes of graduates with loans, particularly at the beginning of their career. Second, outstanding student debt balances are discounted from the amount of resources entrepreneurs can pledge to rent capital. Since entrepreneurial productivity is stochastic and not pre-determined at the time of college enrollment, potentially productive graduates may be later prevented from acquiring capital due to their student debt. By tightening their borrowing constraint, education loans ex-ante reduce and make entry into entrepreneurship more selective, and ex-post limit the expansion of firms run by indebted graduates.

Calibrated to US data, my quantitative framework replicates as untargeted moments several cross-sectional differences between entrepreneurs with and without education, and with or without student debt. I fit closely the share of student borrowers, the business ownership rates of non-college and college graduates, and the composition of the entrepreneurial sample. Moreover, I can replicate between 30 and 80% of the empirically estimated heterogeneity in firm profits, sales and size across owners with and without student debt. Importantly, the model infers sizeable distortions generated by the discounting of education loans from the amount of resources that firm owners can pledge to rent capital. In fact, equalizing access to business credit across individuals with and without student debt would decrease capital misallocation by 4.96%, increase college graduates' entrepreneurial production by 5.39% and result in a 2.11% rise in aggregate output.

The second quantitative contribution of the paper is to use the calibrated framework to analyse the recent rise in student debt and the decline in entrepreneurship for US college graduates. A vast literature has focused on the drop in business dynamism and its potential causes, while few contributions document a steeper decrease in firm ownership rates for college graduates (see Salgado (2020)). I use SCF data to show that a large share of the decline in entrepreneurship for college-educated individuals since the 1980s is driven by graduates with loans. Next, I interpret this finding through the lens of my theory, and, in particular, through the effect of outstanding student loans on entrepreneurial financial constraints. Quantitatively, I compare two different steady states of my model, by varying the return and the price of higher education to match the changes in the college premium and attainment rate between the late 1980s and today. This channel brings about a consistent increase in student debt and in the share of borrowers, and explains a third of the decline in the entrepreneurial rate of US college graduates with loans over the same period.

As an additional exercise, I leverage the exogeneity of the 1998 reform to student debt bankruptcy to establish a stronger empirical link between outstanding college loans and entrepreneurship, which I then replicate in the model. In so doing, I follow a recent strand of literature reviewed in Buera et al. (2021b), which draws upon applied-economics techniques to discipline or validate macro models. In particular, before 1998, college borrowers could discharge their loans after 7 years into repayment, which enables me to exploit the discontinuity in the availability of bankruptcy by repayment year when the reform hit. Employing a regression discontinuity design (RDD) on SCF data, I estimate an elasticity of firm ownership to education loans between 6 and 9%.[4] Then, I use my model to simulate a counterfactual scenario in which some graduates are

[4]This policy change to the provision of college loans bankruptcy is also analysed in Yannelis (2016) in order to link

allowed to discharge their loans after 7 years into repayment, following the legal terms in order before the reform. After matching the share of bankrupt households in the 90s, I show that introducing student debt discharge would increase the entrepreneurial rate of graduates with loans by 8% on average, a partial equilibrium elasticity that closely replicates its empirical counterpart.

Finally, I use the model as a laboratory to study the implications of expanding college grants and borrowing limits, switching to income-based repayment plans and canceling outstanding education loans. All exercises are carried out as neutral to the budget balancing of the government, and allow for GE responses. Note that, while Luo and Mongey (2019) and Abbott et al. (2019) have analysed the effect of changes to college financing on wages and labor market outcomes, there is a lack of understanding with respect to how such policies may impact workers and entrepreneurs differently, and affect aggregate quantities through multiple channels. Specifically, in my model, agents are born with uninsurable heterogeneous wealth and productivity, and student loans foster college enrollment. However, while education ensures higher income growth over the life-cycle, outstanding debt can distort both extensive and intensive entrepreneurial margins. Consistently, higher university subsidies and loans limits, or the possibility of tying debt repayments to one's income all raise college enrollment in my counterfactuals. Moreover, income-driven plans provide relief to indebted graduates when hit by adverse shocks, while merit-based grants can successfully decrease debt overhang for productive students. Yet, changes in the composition of the pool of borrowers and in the average amount borrowed do not increase the entrepreneurial rate of graduates across all policy experiments, and do not necessarily improve capital allocation and output.

Related Literature. This project contributes to a rich macroeconomic literature on financial frictions and entrepreneurship, which has studied the effects of firm borrowing constraints for capital allocation, entrepreneurial decisions and aggregate output (see, among others, Cagetti and De Nardi (2006), Buera et al. (2011) and Midrigan and Xu (2014)). From a theoretical and quantitative point of view, the novel focus of my work is to combine education and entrepreneurial choices together in a heterogeneous agents life-cycle model, which is characterized by the interplay of student debt and its repayment structure with the borrowing constraint faced by entrepreneurs.

Secondly, my work relates to a recent body of applied research that documents several links between student debt and individuals' life choices. For example, Looney and Yannelis (2015a), Yannelis (2016) and Mueller and Yannelis (2019) investigate different trends in repayment and default rates among college borrowers and their potential causes. Parallel to that, Mezza et al. (2020) study the impact of student debt on the likelihood and timing of buying a house, and Di Maggio et al. (2019) show that education loans can cause debt overhang and affect borrowers' geographical mobility and their probability of changing jobs. Catherine and Yannelis (2020) also suggest an effect of college borrowing on family formation. As in Ambrose et al. (2015) and Krishnan and Wang (2019), I concentrate on firm ownership, but use micro-level data to document a relationship between education loans and both the extensive and intensive margins of entrepreneurship. In addition, I complement my findings with a theoretical model and a quantitative exploration.

In combining empirical analyses and a quantitative framework to examine the macroeconomic consequences of student debt, I am similar in spirit to Alon et al. (2021), Ji (2021), Folch and Mazzone (2020), and Luo and Mongey (2019). Differently from these papers, I do not focus on

outstanding student debt and strategic default on other types of credit. More similar to my approach, Krishnan and Wang (2019) study the effect of the 1998 reform on the likelihood of becoming entrepreneurs. There are however key differences with respect to my work: first, I adopt a different regression strategy that exploits the discontinuity in the availability of bankruptcy by repayment year in 1998. Secondly, I use the results to quantitatively validate my model.

human capital accumulation, job search strategies or home-ownership choices, but rather on the consequences of education loans for entrepreneurial outcomes. In this respect, my research relates to a contemporaneous work by Kerdelhué (2021), who explores the impact of college financial aid on entrepreneurship and inequality. Other than using different datasets and empirical strategies, a key distinction of my study is to endogenize the choice of student debt in the model, and investigate the effect of the interplay of education loans and entrepreneurship on capital misallocation and output. Moreover, my quantitative exercises focus on linking the decline in entrepreneurship to the rise in college debt over time, and analysing different policy reforms to college financial aid.

Finally, my paper connects to a growing literature on the macroeconomic effect of higher education policies in the US. For instance, Colas et al. (2021) study the optimal design of student financial aid as a function of parental income, Matsuda (2020) shows that back-loaded subsidies may increase the supply of college-educated labor, and Chatterjee and Ionescu (2012) argue that loan forgiveness for college dropouts could raise welfare. Moreover, Matsuda and Mazur (2022) examine the welfare changes associated with income-contingent plans, while Vardishvili (2020) stresses that decreasing uncertainty over the generosity of financial aid can be welfare-improving. Parallel to that, research by Daruich (2018), Abbott et al. (2019) and Blandin and Herrington (2020) compares government interventions affecting children as opposed to college students, and points out the importance of pre-college investments for university completion and lifetime earnings. With respect to these studies, I investigate how expanding university grants, raising college borrowing limits or switching to income-driven repayment plans affect college enrollment, college graduates' entrepreneurial outcomes, output and welfare. My work is also one of the first assessments of the recent proposal of President Biden to cancel off part of outstanding student loans.

The paper is organized as follows: Section 2 documents the link between student debt and entrepreneurship in SCF data. In Section 3, I develop a life-cycle model of education and occupational choices with college borrowing that is then calibrated to US data in Section 4, where I assess its quantitative fit with respect to my empirical evidence. In Section 5, I study empirically the effect of student debt discharge on entrepreneurial outcomes, and then replicate it quantitatively in the model. In Section 6, I assess the aggregate implications of policy changes to the provision of college financial aid and to the repayment plans of education loans. Finally, Section 7 concludes.

2 Empirical Analysis

In the following section, I present suggestive evidence on the relationship between student debt and entrepreneurship in the US. Specifically, I first focus on the extensive margin of entrepreneurship and show that education loans are associated with a lower likelihood of opening a business. Secondly, I analyse outcomes that regard the intensive margins of entrepreneurship: the presence and extent of college borrowing are linked to a lower probability of receiving business loans, and correlate negatively with business profits, size and revenues. Finally, I discuss whether possible mechanisms of selection into student debt and into entrepreneurship find support in the data.

2.1 Student Debt and Business Ownership

In my empirical exploration, I rely on the SCF, an extensive triennial and cross-sectional survey of US families conducted by the Federal Reserve Board, which provides information on household's demographic characteristics and balance sheet variables, including income, assets and debt.[5] When applicable, it also reports information on respondents' spouses. In my analysis, I use the 1989-2019 dataset and focus on agents in the labor force and between 25 and 65 years old, which leaves me with approximately 170,300 observations.[6] Furthermore, I apply survey weights in regressions and comparative analyses to always ensure the representativeness of my sample.

Even if the SCF does not exclusively target self-employed individuals, these constitute more than 20% of the sample, which makes it suitable for studies of US entrepreneurship (see Michelacci and Schivardi (2020) and Cagetti and De Nardi (2006)). The section related to the businesses owned by respondents contains data on their size, revenues, profits and equity, and information on the 1-digit industry code, the legal status and the funding date of the firms. It also reports how the business was started, the ownership share of the respondents and their working hours. In this paper, I classify as firm owners those that actively manage an enterprise in which they hold the majority share of the ownership, and who report employing at least one salaried worker.[7]

The SCF also asks respondents information regarding their student debt, for example whether they have education loans, the initial amount taken and the amount still to be repaid at the time of the interview, the year in which the loan was issued and started to be repaid, the interest rate charged and the type of repayment plan agents are enrolled into. In the sample period I consider, 20.3% of all respondents affirms to have a student loan to repay,[8] and the average debt taken is around 30,800$, which is in line with estimates from the National Center of Education Statistics.[9]

Table 1: Entrepreneurial Rates: 1989-2019

Educational Level	*Average*	*Without Loan*	*With Loan*
College Graduates	14.23%	15.56%	10.27%
Non College Graduates	10.42%	-	-

As reported in Table 1, entrepreneurial rates of households with college are higher than those of non-college graduates.[10] Yet, firm ownership rates of college graduates with student loans are

[5]Table A1, Table A2 and Table A3 in the Appendix report a list of all the variables used in my regressions, including demographic, business and student debt-related ones, along with a brief explanation and their unit of measure. In Table A4 and Table A5, I instead list several distributional moments and patterns regarding student debt take-up and repayment, comparing the estimates I compute from the SCF to those reported in other available surveys and papers.

[6]In the SCF, the information is stored in five separate imputation replicates (which are denominated as "implicates"). For example, for the 5,783 families interviewed during the 2019 survey, there are 28,915 records in the final data set.

[7]My analysis is robust to considering also self-employed households as entrepreneurs. Moreover, while SCF does not contain information on the reason why entrepreneurs started or operate their business, existing evidence from the Global Entrepreneurship Monitor shows that less than 15% of business owners in the US opens a firm out of necessity.

[8]Before 2016, the SCF did not ask specifically for whose education the loan was taken. For this reason, I first check that my results hold whenever focusing only on the last 2 surveys (2016 and 2019), for which I can identify the correspondence between the person interviewed and the actual underwriter of the student loan reported. Secondly, when pooling together all the survey years, I check that my results hold true whenever focusing on a restricted subsample of the population, namely on those between 25 and 40 years old, which should exclude cases of parents taking or having to repay loans in the name of their children. See Hershbein and Hollenbeck (2015) for discussions on this issue.

[9]Over the last decade in particular, roughly 35% of the US population aged 25 and older is reported to have earned a college degree. Among college graduates, on average 65% have to borrow to finance college. Hence, 23% of the US population above 25 years old is estimated to have negative student loan balances to repay after graduation. Moreover, borrowers on average take out between 30K and 50K $ in student loans, as reported by Looney and Yannelis (2015b).

[10]A large literature has established that education correlates positively with entrepreneurial rates, for example Poschke (2013). College can enhance human capital accumulation, provide individuals and facilitate inter-personal connections or networks that may improve entrepreneurial outcomes. Moreover, starting a firm in particular fields often

substantially lower than for their non-indebted counterpart, and closer to the ones of non-college graduates. Note that Table 1 shows averages for the 1989–2019 period: however, over the recent decades, business ownership rates have decreased and the average amount of student debt per person has increased, as reported in Figure A.1. The first trend speaks to the steady decline in US entrepreneurial rates extensively documented by Decker et al. (2014), while the second one has been argued to reflect changes in the returns to college and in the educational system, for instance regarding tuition costs, loan limits and funding schemes. In Section 4, I will analyse the co-evolution of business rates and student loans over time; here, I instead focus on the cross-sectional differences between individuals with and without college borrowing. To assess the interplay of education loans and firm ownership, I hence estimate the likelihood of becoming an entrepreneur for agents in my sample by running a set of probit regressions of the following form:

$$Pr(BusOwner_{it} = 1) = F\left(\beta_0 + \beta_1 Student\ Loan_{it} + \delta'\mathbf{\Gamma}_i + \gamma'\mathbf{\Phi}_{it} + \alpha_t + \epsilon_{it}\right) \tag{1}$$

where *BusOwner* is a binary variable equal to 1 if individuals are entrepreneurs at the time of the survey, and to 0 if they are not. Focusing on the right-hand side of the equation, three variables can define the explanatory regressor *Student Loan*. First, I use the (log) original amount of student debt taken out by the individual, which does not depend on the survey year. Secondly, I either employ a dummy variable that signals the presence of pending student loans in the balance sheet of the household, or the (log) amount still to be paid as of the interview. Note that 80% of college borrowers has only one recorded loan, while a smaller fraction of the sample reports two to three loans.[11] Here, I consider the total amount of student debt hold or taken out by the households.

Table 2: Business Ownership

	(1)	(2)	(3)	(4)
log(Original Student Debt Taken)	-0.0034***	-0.0024***		
	(0.0002)	(0.0002)		
Dummy(Have Loan)			-0.0241***	
			(0.0024)	
log(Student Debt Still Owed)				-0.0025***
				(0.0003)
Pre-College Controls	Y	Y	Y	Y
General Controls	N	Y	Y	Y
Survey Year FE	N	Y	Y	Y
Observations	170,302	170,302	170,302	170,302
Pseudo-R^2	0.0373	0.0493	0.0521	0.0520

Notes: Estimates are average marginal effects. Robust standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1. Survey weights are used. The dependent variable is a binary indicator = 1 if the individual is a business owner. Pre-College controls refer to agent's gender and ethnicity (Table A8 includes parental education, only available in 2016/2019). General controls are agents' education, age, marital and home-ownership status and income. Robust to including spousal income, households' leverage or assets, and to using either an income or wealth category by age and educational level instead of their personal income. Table A9 provides regression results for the cases in which business ownership is defined without any restriction on the size of the ownership share of respondents.

The regressions include a set of controls $\mathbf{\Gamma}_i$, which capture factors that were pre-determined to the choice of getting a student loan and could affect entry into entrepreneurship (e.g. age, gender, ethnicity and parental education). Finally, I sequentially introduce several control variables recorded at the time of the survey that were not pre-determined to the moment in which individuals contracted college borrowing, such as their education level, marital and home-ownership status,

requires specialized education because of the nature of certain industries (e.e: civil engineers and biologists, etc).

[11] This is the case of separate loans to finance undergraduate and graduate studies, for example.

income or wealth.[12] These latter regressions also include survey year fixed effects (FE).

Table 2 shows that student debt correlates negatively with business ownership. In Column (1), where only pre-determined controls are included, an increase of 1% in student debt is associated with a 0.34% lower likelihood of becoming an entrepreneur. To interpret this result, note that the average entrepreneurial rate in the sample is 12%, while the average size of college borrowing is $31,000. Hence, a $3,100 higher amount of student debt at graduation is associated to a roughly 0.5 percentage points (p.p.) difference in the likelihood of becoming entrepreneurs relative to the mean. In Columns (2)-(4), I then control for variables recorded at the time of the interview that may correlate with business ownership. Both the initial amount of debt taken and the amount still owed at the time of the survey have a negative relationship with business ownership. More specifically, in Column (3), I use as main regressor a dummy variable that is 1 if the respondent reports having a student loan still to repay, and 0 otherwise. In line with the other coefficients, having a student loan is associated with a 3% lower likelihood of becoming an entrepreneur.[13,14]

Table 3: Business Ownership, College Graduates Only

	(1)	(2)	(3)	(4)
log(Original Student Debt Taken)	-0.0058***	-0.0025***		
	(0.0004)	(0.0004)		
Dummy(Have Loan)			-0.0294***	
			(0.0039)	
log(Student Debt Still Owed)				-0.0027***
				(0.0004)
Pre-College Controls	Y	Y	Y	Y
General Controls	N	Y	Y	Y
Survey Year FE	N	Y	Y	Y
Observations	80,157	80,157	80,157	80,157
Pseudo-R^2	0.0202	0.0452	0.0455	0.0453

Notes: Estimates are average marginal effects. Robust standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1. Survey weights are used. The dependent variable is a binary indicator = 1 if the individual is a business owner. Pre-College controls refer to agent's gender and ethnicity (robust to including parental education, only available in 2016/2019). General controls are agents' education, age, marital and home-ownership status and income. Robust to including spousal income, households' leverage or assets, to using either an income or wealth category by age and educational level instead of agents' personal income, and to considering owners with any given equity share.

As a robustness check, Table A10 conducts the analysis focusing only on the first education loan reported by respondents. Moreover, Table 3 estimates again Equation 1 excluding non-college graduates. The presence and extent of student debt correlate negatively with business ownership also among individuals with at least a bachelor degree. Interestingly, the magnitudes of the coefficients across the different specifications are moderately bigger than in Table 2, suggesting that the estimated gap in entrepreneurial rates by student debt is in fact wider within college graduates. This is consistent with the comparison of unconditional entrepreneurial rates shown in Table 1. Moreover, the negative correlation between education loans and entrepreneurship is 3 times stronger when I focus my analysis on the last 15 years of data (i.e: 2005-2019 vs 1989-2004).

[12] Due to the high degree of endogeneity of both income and wealth, I can alternatively use the average income or assets of agents within the same age and/or education category (robust to do it by age and/or education and year too).

[13] Robust to interacting controls with student debt to check results are not driven by one demographic group only.

[14] Results are in line with recent studies on the relevance of available credit and entrepreneurial personal balance sheets for firm ownership and financing (see Herkenhoff et al. (2021) and Robb and Robinson (2014)). It is important to mention that, in a previous contribution, Hurst and Lusardi (2004) found little role for household's net worth in determining entrepreneurship. However, Hurst and Lusardi (2004) focused on self-employment (as opposed to firm ownership) and could not observe instances of financial constraints, as opposed to more recent analyses on the matter.

Next, in Table 4, I investigate whether the link between outstanding student debt and entrepreneurial decisions varies across the distribution of income. For agents with information of personal earnings, I run a set of regressions similar to Equation 1, including the amount of education loans individuals graduated with as an additional control and isolating the effect of their outstanding balances. Column (1) reports results without conditioning or controlling for agents' income: a 1% increase in the amount of student debt owed at the time of the survey is associated with a 0.9% lower likelihood of being an entrepreneur. Yet, the interplay of outstanding education loans and the extensive margin of entrepreneurship is stronger for agents in the bottom half of the earnings distribution, for whom a 1% higher balance on college loans is associated with a 2.6% lower likelihood of being a business owner. Similarly, Table A11 shows that the coefficient attached to outstanding student debt decreases with individuals' age, suggesting a stronger negative correlation between college loans and entrepreneurship early on in agents' life-cycle.

Table 4: Business Ownership and Outstanding Student Balances by Earnings

	Business Ownership			Firms with 20+ Employees		
	Full Sample	Earnings<p50	Earnings>=p50	Full Sample	Earnings<p50	Earnings>=p50
SLoans Owed	-0.0094***	-0.0257***	-0.0071***	-0.0161***	-0.0112***	-0.0507***
	(0.0031)	(0.0059)	(0.0022)	(0.0044)	(0.0028)	(0.0186)
Controls	Y	Y	Y	Y	Y	Y
Survey Year FE	Y	Y	Y	Y	Y	Y
Observations	170,302	83,229	84,287	37,051	23,749	10,521
Pseudo-R^2	0.0381	0.1060	0.0177	0.0263	0.0346	0.0312

Notes: Estimates are average marginal effects. Robust standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1. Survey weights are used. In Columns (1)-(3), the dependent variable is a binary indicator = 1 if the individual is a business owner, and = 0 if not. In Columns (4)-(6), the dependent variable is a binary indicator = 1 if the business owned by the respondent has 20+ employees, and = 0 if it has less than 20 employees. Controls refer to agent's gender and ethnicity, age, student loan size at graduation, educational level, marital and home-ownership status.

When I focus on the likelihood of owning a relatively big enterprise (i.e: 20+ employees), the association between student debt and entrepreneurship becomes stronger. Overall, Column (4) in Table 4 shows that a 1% increase in the amount of student loans owed at the time of the survey is associated with a 1.6% lower likelihood of running relatively big firms. Yet, the effect is larger for individuals in the top half of the earnings distribution (see Column (5)-(6)). Taken together, these results suggest that student debt correlates negatively with entrepreneurship by more for low-income earners, but that education loans may have a stronger influence on business size for top earners. Note also that, when repeating regressions in Columns (3) and (6) for individuals above the 90th percentile of the earnings distribution, I no longer find a significant negative association between entrepreneurship and outstanding student debt, which instead is still negatively correlated with big firms ownership.[15] This may be indicative of the heterogeneity in the barriers to the creation of firms and big-firms along the spectrum of the earnings distribution, and it is consistent with findings in Hurst and Lusardi (2004) and Herkenhoff et al. (2021), who document a non-linear relation between wealth and self-employment, and between credit access and business ownership. Results are robust to doing the analysis by agents' position in the asset distribution.[16]

[15] The coefficient on the amount of student debt owed decreases to -0.0312, significant at the 10% confidence level.

[16] Table A7 reports a robustness check for regressions in Table 2, in which I analyze the correlation between entrepreneurship and education loans controlling for agents' net worth (excluding student debt). This stresses the relevance of student loans for entrepreneurial decisions beyond the effect of individual wealth and additional indebtedness.

Focusing then more closely on business owners and on the enterprises they run, I find that individuals that took out bigger student loans to finance their college education have on average a higher amount of personal wealth collateralized for their businesses, as reported in Table A6. This can suggest that entrepreneurs carrying larger student debt balances might have to provide more collateral to back up their entrepreneurial operations. Along similar lines, Figure A.2 analyses the business legal status of enterprises run by individuals with and without education loans. Indebted owners are less likely to open corporations or limited liability companies, as opposed to individuals without student debt balances to repay.[17] Figure A.2 also shows that college graduates without student loans tend to start their enterprises earlier on in life.[18] In particular, conditional on the same educational attainment, firms of indebted entrepreneurs are 5 years younger, which indicates that households with student debt may delay undertaking their entrepreneurial career.

As pointed out by Alon et al. (2021), the repayment of student loans can incentivize individuals to trade-off higher earnings upon graduation with careers of better long-run prospects. In a similar way, Luo and Mongey (2019) show that agents with student debt generally choose to work for highly-paid jobs with worse amenities early on in their life-cycle, while they are still repaying their loan balances. Consistent with their mechanisms, I argue that college borrowing could similarly discourage or delay business ownership, as opening a firm can lead to potentially higher earnings, but also involves taking higher risk. I will further explore this trade-off in the quantitative section.

2.2 Student Debt and Business Outcomes

After having investigated the link between student debt and the extensive margin of entrepreneurship, I now focus on analysing several business outcomes across entrepreneurs with different education loans balances. In terms of the intensive entrepreneurial margins considered, I first examine business financing, and then turn to business size, sales, profits and profitability measures.

2.2.1 Business Financing

Enterprises generally need funds to run their operations, and one way to obtain finances is through business borrowing.[19] The SCF records information on whether the respondent applied for and obtained a business loan within the 12 previous months before the interview. First, in Table A12, I report estimates for the likelihood of applying for a business loan. Neither the initial amount of student debt taken nor the outstanding balances as of the survey year t correlate with the probability of asking for business funding, suggesting little to no role for any heterogeneity in the demand for credit across indebted and non-indebted entrepreneurs. Secondly, I estimate the likelihood of being turned down in a business loan application via a probit regression of the following form:

$$Pr(LoanApproved_{it} = 1) = F\left(\beta_0 + \beta_1 Student\ Loan_{it} + \delta' \mathbf{\Gamma}_i + \gamma' \mathbf{\Phi}_{it} + \alpha_t + \epsilon_{it}\right) \tag{2}$$

[17] The business legal status reported at the time of the interview is likely to be the one with which the business originally started. Changing the legal status of an enterprise is very infrequent in the US, and bureaucratically complex.

[18] For this comparison, I focus on entrepreneurs that funded their own business, as opposed to inheriting or joining it. In the period considered, 75% of the business owners report funding their own business, 18% buying it, 4% inheriting it, and 3% joining it as a partner. Moreover, more than 95% of entrepreneurs hold a >50% share in their business.

[19] Using a sample of US startups from the Kauffman Firm Survey, Morazzoni and Sy (2021) show that business borrowing from financial institutions tend to represent the most important source of funding for entrepreneurs.

where $LoanApproved$ is a binary variable that takes a value of 1 if the business loan request of entrepreneurs was approved, and 0 if it was rejected. Similar to previous regressions, I first use the (log) initial amount of student debt taken by the individual, then a dummy variable that signals the presence of pending education loans in the balance sheet of the households, and finally the (log) amount still to be repaid at the time of the survey year t. As before, whenever more than one student loan is recorded for a given respondent, I consider the total amount of education debt hold or taken. I introduce individual-level variables and fixed effects as in Equation 1, as well as a set of firm-level controls that include business size, age, legal type and 1-digit industry code.[20]

Table 5: Business Loan Approval

	(1)	(2)	(3)	(4)
log(Original Student Debt Taken)	-0.0120***	-0.0118***		
	(0.0028)	(0.0025)		
Dummy(Have Loan)			-0.1308***	
			(0.0262)	
log(Student Debt Still Owed)				-0.0125***
				(0.0026)
Pre-College Controls	Y	Y	Y	Y
General Controls	N	Y	Y	Y
Firm Controls	N	Y	Y	Y
Survey Year FE	N	Y	Y	Y
Industry FE	N	Y	Y	Y
Observations	5,196	5,075	5,075	5,075
Pseudo-R^2	0.0365	0.2174	0.2188	0.2178

Notes: Estimates are average marginal effects. Robust standard errors in parentheses. ***p$<$0.01, **p$<$0.05, *p$<$0.1. Survey weights are used. The dependent variable is a binary indicator $=1$ if the owner had a business loan application approved within the 12 previous months before the survey interview. Pre-College controls refer to agent's gender and ethnicity (Table A8 includes parental education, only available in 2016/2019). General controls are agents' education, age, marital and home-ownership status and income. Firm controls include size, business age, legal type and individuals working hours. Robust to including spousal income, households' leverage or assets, to using either an income or wealth category by age and educational level instead of agents' personal income, and to considering owners with any given equity share.

As reported in Table 5, student debt shows a negative correlation with business credit approval. In Column (1), I only control for demographic characteristics that were pre-determined at the time the loan was taken, such as gender, ethnicity and parental education: for entrepreneurs, an increase of 1% in student debt is associated with a 1.2% lower likelihood of getting business funding. In all the other specifications, I control for variables recorded at the time of the interview that may correlate with business credit approvals, therefore assessing how student loans of entrepreneurs within the same wealth or business income categories correlate with their likelihood of securing external finances. Column (2) shows that the initial amount of student debt has a 1.3% negative relationship with business loan approval. In Column (3), I instead use as main regressor a dummy variable equal to 1, if the respondent is still in repayment, and to 0 otherwise. Consistent with previous results, having a student loan is associated with a 13% lower likelihood of receiving business credit. Finally, the regression in Column (4) exploits the amount of college borrowing still owed at the time of the survey, finding a similar coefficient to the one in Column (2).[21]

[20] Results are robust to excluding firms offering accounting and legal services, and are available upon request.

[21] I can further restrict the focus only to entrepreneurs that are college graduates. Similar to what observed for entrepreneurial rates in the previous section, this choice reduces noticeably the sample size and gives (statistically significant) stronger effects across the different regression specifications in Table 5. All results are available upon request.

2.2.2 Business Size, Sales, Net Worth and Profits

An impaired access to business financing from external sources is likely to subsequently influence the operations of firms run by indebted entrepreneurs. For this reason, I next focus on the size, sales, net worth and profits of the enterprises in the SCF sample, and examine whether the amount of student debt taken to attend college is associated in any way to these key business performance indicators. For example, due to more severe difficulties in accessing business credit, entrepreneurs that took out larger amounts of education loans or still have to repay substantial balances at the time of the survey interview might be running smaller firms (measured in numbers of employees). Parallel to that, if external credit is used to finance capital acquisition and business operations, firms owned by indebted entrepreneurs might also generate lower revenues and profits, even within the same size category. To test for this hypothesis, I run the following set of regressions:

$$y_{it} = \beta_0 + \beta_1 \textit{Student Loan}_{it} + \delta' \mathbf{\Gamma}_i + \gamma' \mathbf{\Phi}_{it} + \alpha_t + \epsilon_{it} \tag{3}$$

where $y = \{\textit{employees} \, ; \textit{sales} \, ; \textit{profits} \, ; \textit{net worth}\}$ is a vector containing either the number of employees, the (log) gross sales, the (log) profits or (log) net worth of the business as reported by entrepreneurs in the SCF sample at the time the interview took place. The net worth of a business is to be intended as the value at which the business could have been sold in the year of the survey interview. I allow for the sets of firm and individual-level controls previously explained, and include survey year fixed effects (α_t). Results for size and sales are displayed in Table 6, while those concerning business profits and net worth are instead reported in Table A14 and Table A15.

Table 6: Business Outcomes: Size and Gross Sales

	Employees	Employees	Sales	Sales
log(Original Student Debt Taken)	-1.9828***	-1.9919***	-0.0648***	-0.0423***
	(0.1656)	(0.1890)	(0.0055)	(0.0048)
Pre-College Controls	Y	Y	Y	Y
General Controls	N	Y	N	Y
Firm Controls	N	Y	N	Y
Survey Year FE	Y	Y	Y	Y
Industry FE	N	Y	N	Y
Observations	40,145	39,461	37,540	36,855
R^2	0.0026	0.0339	0.0780	0.4054

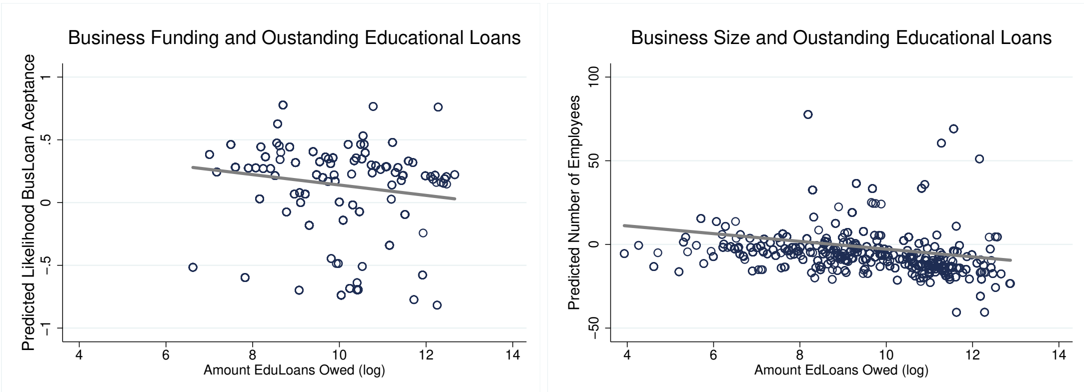
Notes: Robust standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1. Survey weights are used. The dependent variables are either the number of employees or the $\log(Sales)$ of entrepreneurs in the sample. Pre-College controls refer to agent's gender and ethnicity (robust to include parental education, only available in 2016/2019). General controls are agents' education, age, marital and home-ownership status and income. Firm controls include business age, legal type and individuals working hours (and business size in Columns (3)-(4)). Robust to including spousal income, households' leverage or assets, to using either an income or wealth category by age and educational level instead of personal income, and to considering owners with any given equity share.

The estimation of Equation 3 reveals that the amount of student debt taken by entrepreneurs to finance education is linked to a lower business size:[22] specifically, an increase of 1000$ in college borrowing is associated to hiring 12 employees less. Moreover, within businesses of comparable profile (including size) and for entrepreneurs of similar demographic and financial characteristics, an increase of 1% in the initial amount of student debt upon graduation correlates with 4-6% lower sales, 2-4% lower profits, and 5-7% lower business net worth. Results are stable to the sequen-

[22] Table A13 reports robustness checks using a dummy for whether the entrepreneur has a student loan to repay at the time of the interview, as well as the actual amount still to be repaid. Results are in line with the baseline specification.

tial introduction of controls, which suggests that the magnitude and statistical significance of the coefficients of interest are not simply driven by the choice of the regressors included. Also, the coefficient β_1 is almost 3 times larger when considering only firms that are active in Manufacturing and Wholesale Trade industries, two sectors where entrepreneurs tend to need larger finances to operate. Finally, Figure 1 further confirms that, conditional on having taken out loans to finance a degree, the amount still owed at the time of the interview has a negative impact both on the likelihood of having a business funding application accepted and on the size of the business.

Figure 1: Business Outcomes by Student Debt Outstanding Balances



Notes: Residuals from OLS and Linear Probability regressions. Survey weights are used. The dependent variables are the number of employees and a dummy for whether a business loan application was accepted in the 12 months before the survey date (in the panel on the left, the upper limit is set at the p95 of the distribution of residuals for illustration purposes). Control variables are as in the baseline regressions, namely as in Equation 2 and Equation 3.

2.3 Selection Effects

2.3.1 Selection into Entrepreneurship

So far, I have shown that the amount of debt contracted to finance college education, as well as the amount still to repay at a given time throughout individuals' life-cycle, are correlated with a lower likelihood of being an entrepreneur and obtaining funding, and are associated with opening firms of smaller size, profits and sales. To rationalize my results, this paper advances the hypothesis that the financial burden implied by education loans, combined with the risk involved in running a firm, could act as a deterrent to entrepreneurial entry and as a barrier to business activities.[23]

On the one hand, making debt repayments and disburse interest rates for 10 to 25 years after graduation can have a negative income effect on households' available resources, and slow down wealth accumulation, with the effect being relatively more severe for low-income earners. Note that the current average monthly payment college graduates have to make on their student debt is around \$450, nearly 10% of an average monthly salary.[24] Loan repayments may therefore reduce savings and discourage or delay firm ownership, as entrepreneurs' personal asset are crucial for establishing and running businesses (see Quadrini (2009) for a review of the literature and

[23] As mentioned before in the introduction, the concern that student debt could prevent entrepreneurial activities has long been debated by the public opinion. As a nice summary, the magazine Forbes has recently reported the insights of 13 members of its Business Council with respect to how student loan debt can affect an entrepreneur's journey.

[24] See `https://educationdata.org/average-student-loan-payment` and Avery and Turner (2012). Importantly, for agents with graduate loans, monthly payments are on average between \$700 and \$1500. Hua (2021) also documents that the fraction of college graduates with negative net worth has increased over the past decades.

Robb and Robinson (2014) for recent empirical evidence). Consistent with this mechanism, I have shown in Table 4 that, conditional on demographics and on the initial amount of debt taken for college purposes, the negative correlation between outstanding education loans balances and business ownership is three times stronger for individuals below the median in the earnings distribution.

On the other hand, the influence of college borrowing on entrepreneurial outcomes could also be of a financial nature. Lending institutions are known to discount the amount of outstanding debt individuals carry whenever they apply for other types of loans, and the negative effect of student debt on the access to credit is estimated to be more severe in tightly underwritten markets (see Mezza et al. (2021)). Similarly, there is extensive evidence on how student loans are associated not only with a higher likelihood of being credit constrained, (see Folch and Mazzone (2020) and Mezza et al. (2020) on the effect on home-ownership rates) but also with a higher likelihood of declaring consumer credit bankruptcy (see Gicheva and Thompson (2015))[25]. At the same time, Brown et al. (2015) have shown that student debt borrowers have lower average credit scores nowadays as opposed to the beginning of the century, when their risk profiles were comparable to non-borrowers.[26] Consistent with this set of results, it is possible that education loans may also decrease the likelihood of getting funds (or the amount one can get) for running or starting a business,[27] which is in line with the evidence from the SCF data I presented in Table 5.

Interestingly, an impaired access to external credit could also imply that barriers to business ownership may be more pronounced for individuals with higher initial or outstanding college loans balances, resulting in a stronger selection into the entrepreneurial sample. As a consequence, one may also expect individuals who took out or carry large amounts of student debt and became entrepreneurs to be marginally more productive. To check for instances of selection into entrepreneurship, I compute profitability indicators such as profits per dollar revenues, or profits per dollar of collateralized debt. Then, I assess how these profitability measures correlate with student loans for the entrepreneurs in my sample by running the following set of regressions:

$$y_{it} = \beta_0 + \beta_1 \textit{Student Loan}_{it} \times \textit{Business Size}_{it} + \delta' \mathbf{\Gamma}_i + \gamma' \mathbf{\Phi}_{it} + \alpha_t + \epsilon_{it} \tag{4}$$

where $y = \{\log\left(\frac{Profits}{Revenues}\right); \log\left(\frac{Profits}{CollDebt}\right)\}$ is a vector containing the (log) measures of profits per dollar revenues and profits per dollar of collateralized debt based on the information reported by SCF entrepreneurs at the time the interview took place. I regress both variables against the (log) amount of student loans took by the respondent, the size of their business and an interaction term to ensure that results are not driven by bigger (or smaller) firms only. Control variables and survey year FE are as in Equation 3, and results are displayed below in Table 7. Moreover, Table A16 shows the outcomes of robustness analyses that use as main regressors either the (log) amount of college debt still owed at the time of the interview, or a dummy variable that is equal to 1 if the respondent reports pending student loans on their balance sheets, and to 0 otherwise.

Column (1) and (3) report regression outcomes for the simplest specifications, which include the

[25] This seems not to have been the case before the last two decades: Brown et al. (2015) document that the association between student loans and other debt, such as mortgages, credit cards and auto loans used to be positive. This suggests that, historically, student debt might have been an indicator of borrowers having a higher level of education and projected income, a signal of financial prosperity that could have partially changed in more recent days.

[26] Brown et al. (2015) also documents that, amid the general trend of household deleveraging after the financial crisis, debt balances fell much more for borrowers with education loans, especially those with high levels of student debt.

[27] Business loans from financial institutions are the main source or funding for entrepreneurs in the US. Using the Kauffman Firm Survey, Morazzoni and Sy (2021) shows that funding from secondary sources such as family and friends tend to represent less than 8% of outside debt and less than 4% of equity contributions for typical US entrepreneurs.

Table 7: Business Outcomes: Profitability

	$\log\left(\frac{Profits}{Revenues}\right)$	$\log\left(\frac{Profits}{Revenues}\right)$	$\log\left(\frac{Profits}{CollDebt}\right)$	$\log\left(\frac{Profits}{CollDebt}\right)$
log(Original Student Debt Taken)	0.0204***	0.0122***	0.0052**	0.0058**
	(0.0026)	(0.1890)	(0.0019)	(0.0017)
Pre-College Controls	Y	Y	Y	Y
General Controls	N	Y	N	Y
Firm Controls	N	Y	N	Y
Survey Year FE	Y	Y	Y	Y
Industry FE	N	Y	N	Y
Observations	40,150	39,461	40,150	39,461
R^2	0.0230	0.1411	0.0083	0.0575

Notes: Estimates are average marginal effects. Robust standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1. Survey weights are used. Pre-College controls refer to agent's gender and ethnicity (robust to include parental education, only available in 2016/2019). General controls are agents' education, age, marital and home-ownership status and income. Firm controls include business age, legal type and individuals working hours. Robust to including spousal income, households' leverage or assets, to using either an income or wealth category by age and educational level instead of personal income, and to considering owners with any given equity share.

initial amount of education loans (without the interaction term *Student Loan* $\times$ *Business Size*) and controls that were pre-determined at the time the student loan was taken. Columns (2) and (4) consider instead the full set of controls and the interaction term in Equation 4, and show that the coefficient on *Student Loan* is significant and positive, above and beyond confounding effects coming from the size and characteristics of the businesses run by the entrepreneurs. Larger initial amounts of education loans correlate with higher firm profitability per dollar revenues or per dollar of collateralized debt. In particular, a 1% increase in the original amount of college borrowing is associated with 4 to 9% higher business profit margins, which suggests that owners with larger balances of student loans may have also undergone a stricter selection into entrepreneurship.

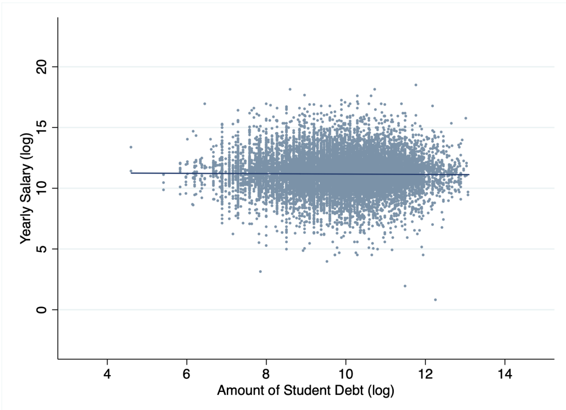
2.3.2 Selection into Student Loans

An important question to ask at this point is who are the individuals that borrow for college education. If agents were to select into student debt on characteristics that are also linked to a lower likelihood of becoming entrepreneurs and running productive firms, my results would primarily capture such selection mechanism and have little to say about the financial or income effects of education loans on entrepreneurial outcomes. I tackle this issue in two steps: in the remainder of this section, I first discuss suggestive evidence of the fact that my results do not seem to be biased by major instances of negative selection into college borrowing. Then, in Section 5, I exploit an exogenous change in the repayment policy of student debt to strengthen the correlation between education loans and entrepreneurship, above and beyond confounding selection effects.

What does it take to open and run a firm? Data and theories converge on pointing at wealth and entrepreneurial ability as two crucial factors behind the majority of business stories (see Cagetti and De Nardi (2006) and Buera et al. (2011) for example). Hence, it seems important to assess whether student debt borrowers significantly differ from non-borrowers along these two dimensions. Let's first focus on family and personal wealth. Parents' finances have been and still are an important determinant of college attendance for US high school graduates. However, family wealth has recently become a weaker predictor for the likelihood of borrowing to finance education (see Lochner and Monge-Naranjo (2016)). According to estimates from the National Center for Education Statistics, student loans growth at the extensive margin (percent borrowing) and at the intensive margin (amount per borrower) was actually more pronounced for the highest family income quartile over the 1989-2004 period (see Berkner (2000) and Wei and Berkner (2008)). Federal data also shows that, among graduates from the 2012 cohort, half of the students from higher-income families borrowed for college, twice the share compared to the 1992-93 cohort.[28]

Such steady increase in the share of borrowers from the highest family income quartiles might reflect the introduction of unsubsidized federal loans, which can be taken out irrespective of financial need,[29] and the fact that other aid schemes such as Pell Grants only target low and middle-income students that qualify for subsidies.[30] Along these lines, Looney and Yannelis (2015b) have shown that rich US households are now more and more likely to use education loans to pay for tuition and boarding costs, especially at top universities and Ivy League Schools.[31] Although I have no data on family wealth or income for the SCF respondents, I have controlled for parental education in my baseline regressions, and interacted it with college debt as well. In addition, individuals with student loans are those who acquired a higher education and hence will likely have better career prospects and earnings profiles.[32] Consistently, I have also interacted individuals' student debt balances with their wealth or income percentile, without registering changes to my main results. Therefore, this preliminary exploration does not suggest that the link between student loans and entrepreneurship is due to the fact that borrowers only pertain to the bottom of the income or wealth distributions, considering either their family's or their personal finances.

Figure 2: Yearly Salary and Student Debt Balance upon Graduation



Parallel to that, it seems difficult to argue that student loan borrowers might clearly and significantly have lower entrepreneurial skills compared to non-borrowers. Entrepreneurial skills are complex to measure but are typically proxied by educational attainments (see Poschke (2013)). Only students that choose to acquire higher education get a student loan, and higher education has been often found to positively impact entrepreneurial outcomes (see Guo et al. (2016) and Michelacci and Schivardi (2020)). Comparing individuals with and without student debt, one should in fact expect the latter group to be more likely to open and run a successful business.

Shifting the focus to college graduates, a possible confounding mechanism would be that students

[28]See www.pewresearch.org/social-trends/2014/10/07/.

[29]Unsubsidized Federal loans were first introduced with the 1992 Higher Education Act.

[30]In the case of Pell Grants, the typical share of college tuition covered is less than 30%. See: www.brookings.edu/research/the-economic-case-for-doubling-the-pell-grant/

[31]www.brookings.edu/opinions/students-at-elite-schools/

[32]www.peoplespolicyproject.org/2020/11/16/what-is-the-current-student-debt-situation/.

with higher talent get more often access to grants and hence do not borrow for their degrees. Merit-based aid is overall limited and infrequent, offsetting at most 20% of the average financial needs of relatively few perspective students (roughly 15% of them), and leaving the rest to be covered either by family contributions or through borrowing. More importantly, the negative selection of students into borrowing finds little support empirically. To provide a finer measure for individuals' talent, I use data from the US National Longitudinal Survey of Youth (NLSY97), which contains results to cognitive and attitudinal tests administered to all respondents, irrespective of their educational level. In Table A17, I show that cognitive abilities are a strong predictor of both receiving grants and taking out student debt for college. Results also illustrate that cognitive abilities do not correlate with getting higher amounts of grants as opposed to education loans.

Secondly, if students were to negatively select into having education loans according to their idiosyncratic skills, it should be reflected in their entrepreneurial outcomes. Contrary to that, I have shown in Table 7 that profitability measures are higher (not lower) for college borrowers. Using NLSY97, in Table A18 I then regress self-employment rates on the interaction between individuals' education loans and cognitive abilities to show that student debt per sé is still significantly and negatively associated with entrepreneurial outcomes. Along this line, both Luo and Mongey (2019) and Alon et al. (2021) have documented that regressing individuals' wages on student debt leads to non-significant correlations, which I can also verify in the SCF data (see Figure 2).[33] In this sense, I can rule out that individuals with student loans are evidently the least productive ones among college graduates, otherwise it should likely be reflected in their earnings as well.

To conclude the discussion, I check for any heterogeneity in the correlation between college borrowing and entry into entrepreneurship by student debt balances. Specifically, I estimate the elasticity of business ownership to outstanding college debt – analogous to β_1 in Equation 1 and in Column (4) of Table 2 – for individuals in different quartiles of the student loans distribution. As shown in Figure A.4, conditioning on the respondent having taken out college debt, the negative correlation between outstanding balances and the extensive margin of entrepreneurship is stronger for agents below the 75^{th} percentile in the student debt distribution. This is in principle consistent with the fact that individuals with larger education loans (above the 75^{th} percentile of the distribution, for example) might be those who attended costly and possibly prestigious universities. The fact that, for these respondents, education loans do not correlate with entrepreneurial entry suggests that they could be positively selected into having student debt either on their wealth, their talent or both. This could further confirm that my results do not depend on respondents with large education loans, and that having large education loans does not necessarily mean being negatively selected on characteristics that predict worse entrepreneurial outcomes.

Using SCF data has still several limitations, above and beyond not observing individuals' idiosyncratic productivity or their family's income. The dataset lacks information on the college fees paid by respondents, as well as on the amount of grants and family contributions they have benefited from. Since student debt and occupational choices are endogenous outcomes that reflect the selection of individuals along their main characteristics, I next build a model that can account for both decision margins and their interaction. Then, in Section 5, I will go back to the data and exploit an exogenous policy change to the availability of education loans bankruptcy in order to reinforce the idea that student debt can indeed have a negative effect on entrepreneurial outcomes.

[33] Their IV set ups also confirm that education loans have a positive effect on wages upon graduation.

3 Model

This section presents a general equilibrium framework that nests together education decisions and occupational choices over the life-cycle of individuals. Households in the model live through three main stages: education, working life, and retirement. They are born with heterogeneous wealth and idiosyncratic productivity, which accumulate and change over time. During youth, agents decide whether to attend college by paying a tuition (net of grants), and whether to take out student loans, which are repaid after graduation. Households are endowed with one unit of time that they either supply inelastically, if they choose to be workers, or use to run a firm, if they choose to be entrepreneurs. They save out of their income and consume a final good, which is produced by the entrepreneurial sector. In particular, output is obtained combining productivity, capital and labor, and entrepreneurs face a limited pledgeability constraint to rent capital.[34]

In the model, student debt and entrepreneurial choices are interconnected because of two main channels: first, loan repayment upon graduation reduces the amount of available resources that individuals can save, and slows down the accumulation of wealth. Since personal assets are the collateral against which entrepreneurs borrow to finance capital acquisition, this mechanism has a direct negative effect on the entrepreneurial rates and outcomes of college graduates with student loans, particularly at the beginning of their career. Secondly, during the repayment period, borrowers' outstanding balances are discounted from the amount of personal resources that can be pledged to finance capital acquisition. By tightening their collateral constraint, student loans ex-ante reduce entry into entrepreneurship, and ex-post limit the expansion of firms run by indebted college graduates. The model can hence account for the interplay of student debt with both the extensive and intensive margins of entrepreneurship, as I have documented in Section 2.

3.1 Primitives and General Settings

Preferences: Agents have a strictly increasing and concave utility function over consumption, which satisfies standard Inada conditions, and whose coefficient of risk aversion is denoted by γ:

$$u(c) = \frac{c^{1-\gamma} - 1}{1-\gamma}$$

Moreover, individuals discount their utility over future consumption at rate β.

Timing: Households are born as if they were out of high-school. In the first stage of their lives, T_{edu}, they decide whether to attend college or to enter directly the labor force. In the years between $T_{edu} + 1$ and T_{work}, all agents work, consume and save. Between $T_{work} + 1$ and T_{end} they retire and live off their savings and pensions until death. Survival probabilities vary by age and are denoted by θ_{age}. To ease the exposition, I suppress time subscripts whenever they are not strictly necessary.

Productivity: Individuals are characterized by heterogeneous idiosyncratic entrepreneurial productivities z, which evolve stochastically over time according to a standard AR(1) process:

$$z' = \rho z + \epsilon \quad \text{with} \quad \epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$$

[34]In the Appendix, I also consider a version of the model augmented with an unconstrained productive sector.

Such process is defined by a conditional distribution $d\Xi(z'|z)$, where I indicate by ρ_z the persistence of individuals' productivity and by ϵ its idiosyncratic risk component. Both terms do not vary by college attainment. Note that my theoretical framework features idiosyncratic shocks to productivity but no source of aggregate uncertainty. As in models à la Buera et al. (2011), I assume that z evolves constantly in the background of individuals and over their life-cycle, regardless of their occupational choices. However, this idiosyncratic productivity component is used only in entrepreneurial production, and does not scale the wage of workers. This is why z can be thought as agents' comparative advantage of pursuing entrepreneurship as opposed to paid work.

Parallel to that, all households – including workers and entrepreneurs – are also characterized by a tenure or efficiency profile, denoted by ℓ^i_{age}, which differs by educational attainment i and evolves exogenously and deterministically over the life-cycle according to the following process:

$$\ell^i_{age} = \zeta^i_1 \times age - \zeta^i_2 \times age^2 \quad \text{with} \quad i \in \{college, nocollege\}$$

Note that the parameters ζ^i_1 and ζ^i_2 govern the slope and curvature of the deterministic efficiency profile of individuals, and will reflect heterogeneities in the income growth of households across educational attainments and over their life-cycle. In this modeling choice, I hence embed the college premium that determines the incentives of young adults to acquire a university degree.

It is important to stress that total entrepreneurial productivity comprises both a deterministic and a stochastic or idiosyncratic component, given by the expressions for ℓ^i_{age} and z previously discussed.[35] Total entrepreneurial productivity is then given by their combination according to:

$$e^{\xi^i_{age}} = e^{\ell^i_{age}} \times e^z \quad \text{with} \quad i \in \{college, nocollege\}$$

Firm's Technology: Output is produced through a standard production function that combines total entrepreneurial productivity ξ^i_{age}, capital k and labor l. The production function is increasing in all its arguments, strictly concave in capital and labor, and decreasing returns to scale, allowing for a non-degenerate distribution of the enterprise size. In particular, $f(\xi^i_{age}, k, l)$ is given by:

$$f(\xi^i_{age}, k, l) = e^{\xi^i_{age}}(k^\alpha l^{1-\alpha})^{1-\nu}, \quad \text{with} \qquad 0 < 1-\nu < 1, \quad \text{and} \quad i \in \{college, nocollege\}$$

where α is the capital share in production and $1-\nu$ is the span of control as in Lucas (1978). Both capital and labor are static inputs and rented on their respective markets at each point in time.

Financial Markets: There is a perfectly competitive intermediary sector that receives deposits from savers and lends funds to firms, without intermediation costs. The rental rate of capital is given by r, the deposit rate which is determined in general equilibrium. Financial markets are incomplete, and entrepreneurs can borrow intra-temporally up to a fraction of their assets a, net of any education loan d they might carry at a given time t. Capital constraints are hence given by:

$$k \leq \lambda(a - \eta d); \qquad a \geq 0$$

where $a \geq 0$ (intertemporal borrowing is ruled out for simplicity) and λ measures the degree of the financial constraint, as micro-founded in Buera et al. (2011). If $\lambda = 1$, agents operate in a zero credit environment, as opposed to the case in which $\lambda = \infty$ and individuals can borrow

[35] As argued by Michelacci and Schivardi (2020), entrepreneurs with graduate degrees have higher returns to their business ventures, and increasingly so by previous experience. This rationalizes including the deterministic life-cycle efficiency component – and the college premium embedded into it – within overall entrepreneurial productivities.

according to their productivity, regardless of their (net) financial wealth. Importantly, the presence of education loans in the balance sheet of college graduates limits the amount of collateral they can pledge to rent capital on financial markets at any given time t during repayment. The parameter η governs precisely the extent to which college loans reduce the amount of wealth entrepreneurs can use to back up capital renting.[36] Note that $d = 0\ \forall t$ in the case of entrepreneurs without college education and for entrepreneurs with college education that did not take out student debt. Moreover, d becomes 0 when indebted college graduates finish repaying their student loans.

Profit Maximization: Entrepreneurs' profit maximizing problem in a given t reads as follows:

$$\pi^* = \max_{l,k} \left\{ e^{\xi^i_{age}} (k^\alpha l^{1-\alpha})^{1-\nu} - wl - (r+\delta)k, \quad \text{s.t.} \quad k \leq \lambda(a - \eta d) \right\} \tag{5}$$

where the price of output is normalized to 1. All entrepreneurs pay capital rental costs $(r+\delta)k$ and salaries wl as variable input costs, where I denote by δ the depreciation rate of capital. Importantly, in this baseline version, I abstract from any other type of production costs, including fixed ones.[37] Moreover, the differences in the profit maximization problem of individuals with and without college education are given by the different processes that characterize their idiosyncratic total entrepreneurial productivity ξ^i_{age}, and by the capital constraint, which varies according to the presence of student loans in the balance sheet of the households. There is no further source of heterogeneity by education in the production technology or in the input prices paid by entrepreneurs.

Occupation Choice: In each year during their working life and until retirement, agents decide their occupation o, based on their wealth a, idiosyncratic comparative advantage as entrepreneurs z, and on the amount of outstanding student debt d. Households choose to be either entrepreneurs ($entr$) or workers ($work$).[38] Entrepreneurs own a firm and earn business profits $\pi(a, z, d, age; r, w)$, while workers inelastically supply one unit of labor and receive an efficiency-

[36] This formulation for the collateral constraint faced by entrepreneurs builds on the way financial frictions are modeled in Buera and Shin (2013), and can be related to the microfoundations of collateral pledgeability limits by Kiyotaki and Moore (1997). In particular, I assume that the amount of debt entrepreneurs can take intra-temporally to finance their operations cannot exceed the returns on capital (which I call ϕk), and that creditors discount student loans from the business debt they grant to owners. Note that the model does not allow for assets to be negative, and creditors are concerned about the amount of liabilities that have to be served in a given period. Since student loans cannot be discharged in bankruptcy (i.e: they are a senior form of debt) , I assume that creditors do not discount from the amount of pledgeable assets all outstanding student debt, but rather a fraction ηd. This latter term reflects the sum of per-period principal payments and interest rates that may be due intra-temporally. Denoting firm liabilities by b one can write:

$$b + \eta * d \leq \phi k$$

Since the net wealth entrepreneurs carry on to the next period is $a = k - b$, I rewrite the borrowing constraint as:

$$k \leq \frac{1}{1-\phi}(a - \eta d) \quad \rightarrow \quad k \leq \lambda(a - \eta d)$$

where I denote by λ the borrowing multiplier given by $\frac{1}{1-\phi}$. For example, let us focus on the zero-credit-environment case in which $\lambda = 1$: what I assume is that, due to the risky nature of running a firm, college-indebted entrepreneurs are required to keep a buffer to face the periodic repayment on student debt they owe to federal authorities.

[37] Including a fixed cost of operation as in Buera and Shin (2013) would strengthen selection into entrepreneurship, by introducing a non-convexity and making a given technology feasible only if operated above a minimum scale. Moreover, stochastic fixed costs that are realized at the end of any given period t would increase the risk associated with opening a firm. This would in turn affect entry patterns and amplify misallocation, but would not change the qualitative implications of the effect of student loans on entrepreneurial margins. I hence consider my baseline version to be a lower bound for the potential distortions created by student debt overhang on business entry and operations.

[38] I abstract from any inter-generational transmission of businesses because inheritances make up for only 4% of existing enterprises in the SCF over the three decades considered. This result from the SCF compares well to other surveys: for example, Kaplan and Rauh (2013) find that more than 80% of Forbes 400 businesses in 2011 were first-generation.

units salary $\tilde{w}_{age}$, given by the general equilibrium wage w and scaled according to the their age-dependent efficiency profile: $\tilde{w}_{age} = e^{\ell^i_{age}} * w$. Recall that, in this baseline version of the model, I have assumed wages to be fully deterministic, while entrepreneurial profits have an uncertainty component.[39]

3.2 Educational Period

Agents start their life with heterogeneous wealth a and heterogeneous idiosyncratic productivity z. The distribution of initial assets and productivity in the economy is stationary and denoted by $F(a,z)$, whose parametrization will be characterized quantitatively in the calibration exercise. Moreover, I assume that initial assets – interpreted as parental wealth – and individual productivity are correlated at birth, to reflect well-documented evidence on the inter-generational persistence of wealth and labor market outcomes in the US. Even though the model does not feature overlapping generations with altruism and/or paternalism, the relation of a and z at birth can be calibrated to deliver the correlation in earnings across generations reported in Chetty et al. (2014).

Young households have to make an education choice and decide whether to attend university or not. College entails a tuition χ, net of subsidies s funded by the government, which are both proportional to individuals' idiosyncratic productivity (i.e: *merit based*) and inversely related to individuals' wealth (i.e: *means-tested*),[40] as it will be further explained in the calibration exercise. College tuition can be paid also by contracting student debt, denoted by d, which is administered by the government.[41] Since applying for financial aid in the US is free and done on-line through the Free Application for Federal Student Aid (FAFSA) form,[42] I do not include costs for obtaining grants or federal loans. Note that young households are heterogeneous in their initial financial resources, and there are no markets to insure against being unable to pay for university. Since college is costly, education grants and debt therefore facilitate enrollment into higher education.

Since agents spend 4 years in college, corresponding to 1 stage in the model, I report the correct time conversion in the Appendix. To ease notation, here V^c and V^{nc} refer to agents' value function during youth – based on their education decision – while W^c and W^{nc} are the continuation values

Moreover, Hurst and Lusardi (2004) reports that 6.5% of business owners covered by the 1993 National Survey of Small Business Finances inherited their firm, while the analysis of the 1992 US Census survey on the Characteristics of Business Owners by Fairlie and Robb (2009) reveals that roughly 2% of owners inherit their business.

[39] In this respect, I follow typical assumptions in standard entrepreneurship models, such as those developed in Buera et al. (2011) and Midrigan and Xu (2014), who assume stochastic idiosyncratic productivity to matter only for entrepreneurs, and focus on the relative risk of opening a firm as opposed to be a worker. To stress the plausibility of this modeling choice, I use SCF data and further show in Figure A.3 that, while the average wage and profit of individuals tend to diverge over their life-cycle, measures of relative volatilities (and hence risk) stay the same.

[40] Nowadays, 80% of students with family income below 30K$ receive a Pell grant, and the award does not generally vary by college. Pell grants can fund enrollment at accredited institutions, including 2-years or part-time programs.

[41] In the US, 90% of college borrowers – worth 92.6% of the total value of student debt – receive loans from Federal Sources. The remaining 10% of the students who borrow for their degrees obtains credit from private lenders, which however have different borrowing conditions. For a perspective on this topic, see the study of Ionescu and Simpson (2016), who quantitatively assess the macro effects of the private market for student loans on college enrollment.

[42] After prospective students submit their FAFSA form, the US Department of Education computes the expected family contribution based on students' dependency status, family size and income. Then, the financial offices of universities put together aid packages for incoming students before the start of the term. Under the Federal Title IV Aid program established in 1965, financial aid is generally offered in the form of loans, grants and, sometimes, work-study plans.

during their working stage. The maximization problem for agents that decide to go to college is:

$$V^c(a,z,age) = \max_{a',d_{edu},c} \left\{ u(c) + \beta\theta_{age} \int W^c(a',z',d',age')d\Xi(z'|z) \right\}$$
$$\text{s.t.} : c + a' = (1+r)a - \chi + d_{edu}$$
$$\text{and} : a' \geq 0, \quad c \geq 0, \quad 0 \leq d_{edu} \leq \underline{d}$$

where $\underline{d}$ is the student debt borrowing limit, which will be pinned down numerically in the calibration based on the average maximum amount of education loans granted for college.[43] Note that federal loans borrowing limits in the US are such that students cannot borrow up to the entire amount of the tuition, which means that $\underline{d} < \chi$. This is a source of social inefficiency in the model economy, as the private cost of higher education exceeds its social cost: in Section 6, I will analyze the gains that can be obtained by increasing borrowing limits and expanding grants provision.

Agents that do not go to college enter directly the labor market, make an occupational choice and decide whether to work for a salary or to become entrepreneurs and earn the net profits generated by their firm.[44] Their value function during youth, denoted by V^{nc}, is given by:

$$V^{nc}(a,z,age) = \max \left\{ V^{nc,work}(a,z,age), V^{nc,entr}(a,z,age) \right\} \tag{6}$$

which accounts for the occupational choice made by non-college-educated individuals that decide to enter directly the labor market. More specifically, the problem for workers can be expressed as:

$$V^{nc,work}(a,z,age) = \max_{a',c} \left\{ u(c) + \beta\theta_{age} \int W^{nc}(a',z',age')d\Xi(z'|z) \right\}$$
$$\text{s.t.} : c + a' = (1+r)a + (1-\tau)\tilde{w}_{age}$$
$$\text{and} : a' \geq 0, \quad c \geq 0$$

where τ denotes the income tax levied by the government. The value function of agents that choose entrepreneurship as their occupation is instead summarized by the following expression:

$$V^{nc,entrep}(a,z,age) = \max_{a',c} \left\{ u(c) + \beta\theta_{age} \int W^{nc}(a',z',age')d\Xi(z'|z) \right\}$$
$$\text{s.t.} : c + a' = (1+r)a + (1-\tau)\pi(a,z,age;r,w)$$
$$\text{and} : a' \geq 0, \quad c \geq 0, \quad k \leq \lambda a$$

Finally, the education choice made by young households boils down to the following decision:

$$\max\{V^c; V^{nc}\}$$

namely to comparing the present and continuation value of getting or not a college degree. Note

[43] In the US, eligibility for federal student loans (except parent PLUS loans) is universal. Loan limits are more binding for undergraduate borrowers, while graduate students can borrow up to the entire cost of their program. More specifically, at the undergraduate level, loan limits vary across the first, the second and the third/fourth year in college, between two broad categories of family-dependency status and across types of loans (eg: direct subsidized vs direct unsubsidized loans). Since I abstract from dependency statuses or debt types and I model college as a one-period choice (i.e: no further graduate education is considered), I assume everyone face the same limit $\underline{d}$ on student loans.

[44] I am ruling out the possibility for work-study combinations while in college, especially due to the assumption of inelastic labor supply. Currently, 40% of full time students work during university, but the vast majority of them have part-time jobs that total at most 20 hours per week. On the contrary, full time workers in college tend to be both older and householders, and represent 8% of the total (`www.nces.ed.gov/programs/coe/indicator/ssa`).

that the choice to attend university is made once in the model, without the possibility of dropping out of college. While college dropout is certainly an important phenomenon to keep in mind for future extensions of the present paper, Abbott et al. (2019) already argue that it occurs far more often in freshman years and among part-time students. Here instead, I limit my focus to full-time students and college-completers, assuming full commitment to graduating from university.

3.3 Working Period

In each year t between $T_{edu}+1$ and T_{work}, all households make consumption and saving decisions and choose their occupation. For households that attended college, the value function V^c to maximize is defined over agents' assets, productivity, student debt and age and given by:

$$W^c(a,z,d,age) = \max\left\{W^{c,work}(a,z,d,age), W^{c,entr}(a,z,d,age)\right\} \tag{7}$$

which accounts for the occupational choice made by college graduates. More specifically, the value function for college-educated working individuals can be written in the following form:

$$\begin{aligned} W^{c,work}(a,z,d,age) &= \max_{a',c}\left\{u(c) + \beta\theta_{age}\int W^c(a',z',d',age')d\Xi(z'|z)\right\} \\ &\text{s.t.}: c + a' = (1+r)a + (1-\tau)\tilde{w}_{age} - \mathcal{R} \\ &\text{and}: a' \geq 0, \quad c \geq 0, \quad \mathcal{R} = \max\left\{\frac{d_{edu}}{T_{repay}} + r^d d, 0\right\} \end{aligned}$$

where $\mathcal{R}$ is the repayment function of student debt. Note that, in my baseline model, there is no element of uncertainty in the repayment of college borrowing: I will discuss in greater detail the introduction of income-based repayment plans and the implications of education loans bankruptcy in the quantitative section of the paper. In the Appendix, I also consider a modified version of the $\mathcal{R}$ function, which allows for periods of non-repayment through student debt forbearance.

During the repayment period, households with education loans have to pay a fixed amount of the original balance, where d_{edu} denotes the accumulated debt in stage 1 of their life. Moreover, they also need to pay an interest on the outstanding principal amount. Importantly, r^d is calculated including a wedge on top of the overall general equilibrium interest rate r, and it is not allowed to fluctuate over the life of the education loan, which reflects the fact that student debt interest rates have become fixed (as opposed to floating) since 2006.[45] I denote by T_{repay} the established repayment length, which is assumed to be the same for all borrowers, independently of their initial or current balances.[46] The law of motion of outstanding student debt is hence given by:

$$d' = (1+r^d)d - \mathcal{R}$$

[45] The interest rate is computed as a percentage of the unpaid principal amount, and it is set by Federal Laws based on the 10-year treasury note rate of a given year. For example, in 2021, subsidized and unsubsidized loans to undergraduate students carried roughly a 4% interest rate, unsubsidized loans to graduate students had a 5.5% interest rate and parent PLUS loans involved almost a 6.5% interest rate. Between 2006 and 2013, interest rates were much higher on average, oscillating between 5% for subsidized undergraduate loans and 8.5% for parent PLUS loans.

[46] To preserve tractability, an important simplifying assumption of my model is to rule out the possibility that agents make excess repayments on their loan to pay it off more quickly. However, looking at SCF data over the 1989-2019 period, roughly 25% of the student debt borrowers interviewed affirmed to be making payments ahead of schedule.

The value function of college-graduates that choose entrepreneurship can be characterized by:

$$
\begin{aligned}
W^{c,entrep}(a,z,d,age) = \max_{a',c} & \left\{ u(c) + \beta\theta_{age} \int W^{c}(a',z',d',age')d\Xi(z'|z) \right\} \\
\text{s.t.} : & \; c + a' = (1+r)a + (1-\tau)\pi(a,z,d,age;r,w) - \mathcal{R} \\
\text{and} : & \; a' \geq 0, \quad c \geq 0, \\
\text{and} : & \; k \leq \lambda(a - \eta d), \quad \mathcal{R} = \max\left\{ \frac{d_{edu}}{T_{repay}} + r^{d}d, 0 \right\}
\end{aligned}
$$

with college debt d' following the same law of motion outlined above.[47] Having defined the value functions of college graduates, one can clearly see the second important departure from social efficiency that is embedded in the model, namely that private returns to higher education are lower than their social counterpart. This is due to three main factors: first, the economy features incomplete markets, which introduce an element of uncertainty in the returns to educational investments, specifically with respect to entrepreneurial careers. Second, there are firm financial frictions that also depend on the presence and extent of student loans. Finally, distortionary taxation reduces the gains from obtaining a higher efficiency profile through college education. These elements will be also key to understand the effects of the policy reforms analysed in Section 6.

As before, the value function of agents that do not go to college, W^{nc}, is instead given by:

$$
W^{nc}(a,z,age) = \max\left\{ W^{nc,work}(a,z,age), W^{nc,entr}(a,z,age) \right\} \tag{8}
$$

which accounts for the occupational choice made by non-college graduates. More specifically, the value function for working individuals without a university degree has the following form:

$$
\begin{aligned}
W^{nc,work}(a,z,age) = \max_{a',c} & \left\{ u(c) + \beta\theta_{age} \int W^{nc}(a',z',age')d\Xi(z'|z) \right\} \\
\text{s.t.} : & \; c + a' = (1+r)a + (1-\tau)\tilde{w}_{age} \\
\text{and} : & \; a' \geq 0, \quad c \geq 0
\end{aligned}
$$

The value function of non-college graduates that choose entrepreneurship is instead given by:

$$
\begin{aligned}
W^{nc,entrep}(a,z,age) = \max_{a',c} & \left\{ u(c) + \beta\theta_{age} \int W^{nc}(a',z',age')d\Xi(z'|z) \right\} \\
\text{s.t.} : & \; c + a' = (1+r)a + (1-\tau)\pi(a,z,age;r,w) \\
\text{and} : & \; a' \geq 0, \quad c \geq 0, \quad k \leq \lambda a
\end{aligned}
$$

At the end of their working life, households retire. In T_{work}, the continuation value that characterizes agents problem is given by $U^{c}(a, z_{T_{work}}, age)$ and $U^{nc}(a, z_{T_{work}}, age)$, further explained below.

[47] Since agents choose the maximum value between becoming entrepreneurs or workers, the net salary $(1-\tau)\tilde{w}_{age}$ is to be considered as the minimum income they can dispose of with certainty in a given year. This feature of the model excludes the possibility of defaulting on the repayment of student debt, even when individuals have no savings.

3.4 Retirement Period

Between $T_{work}+1$ and T_{end}, households make consumption and saving decisions as retirees. They all receive a pension p^i for $i \in \{college, nocollege\}$, which is funded by the government and represents a given share of the income earned in their last working period. Since agents differ in income profiles according to their education $i \in \{college, nocollege\}$ and entrepreneurial productivity z, pensions vary across individuals with and without a college degree, and are affected by the realization of z in $t = T_{work}$. Note that, both in the last year of their lives T_{end} and throughout their life-cycle, households leave any remaining assets upon death as bequest to the next cohort. This ensures that the wealth distribution of new generations remain stable and can be pinned down quantitatively. For households that attended college, the value function U^c to maximize during retirement is defined over their assets, last-working-period productivity, and age and given by:

$$U^c(a, z_{T_{work}}, age) = \max_{a',c}\{u(c) + \beta\theta_{age}U^c(a', z_{T_{work}}, age')\}$$
$$\text{s.t.} \quad a' = (1+r)a - c + p^c \quad \text{and} \quad a' \geq 0, \quad c \geq 0$$

For households that did not attend college, the value function U^{nc} to maximize is defined over agents' assets, last-working-period productivity, and age and can be characterized as follows:

$$U^{nc}(a, z_{T_{work}}, age) = \max_{a',c}\{u(c) + \beta\theta_{age}U^{nc}(a', z_{T_{work}}, age')\}$$
$$\text{s.t.} \quad a' = (1+r)a - c + p^{nc} \quad \text{and} \quad a' \geq 0, \quad c \geq 0$$

3.5 Government

The role of the government in the model is twofold. On the fiscal side, the public sector collects income taxes (the tax rate has been denoted by τ throughout the exposition) and provides pensions to retired agents. On the education side, the government issues student loans and holds in place grants schemes to foster enrollment in college, especially for low-income households. While both the pension rate and the extent of the grant scheme are calibrated quantitatively to match their empirical counterparts, the tax rate τ is a general equilibrium outcome and has to clear the government budget constraint. In particular, fiscal revenues from tax collection are given by:

$$\sum_{t=T_{edu}}^{T_{work}} \int \tau * (\max\{\pi_t(a,z;r,w); \tilde{w}_t\} dH_t^{nc}(a,z)) + \sum_{t=T_{edu}+1}^{T_{work}} \int \tau * (\max\{\pi_t(a,z,d;r,w); \tilde{w}_t\} dH_t^c(a,z,d))$$

where $H_t^{nc}(a,z)$ and $H_t^c(a,z,d)$ denote the distribution of non-college and college households in each time t. Parallel to that, the items in government expenditure are given by pensions:

$$\sum_{t=T_{work}+1}^{T_{end}} \int p * \tilde{w}_{age T_{work}} dH_t^{nc}(a,z) + \sum_{t=T_{work}+1}^{T_{end}} \int p * \tilde{w}_{age T_{work}} dH_t^c(a,z)$$

and by college loans d_t and grant schemes $s_{T_{edu}}$ according to:

$$\sum_{t=T_{edu}}^{T_{repay}} \int d_t dH_t^c(a,z,d) + \int s_{T_{edu}} * dH_{T_{edu}}^c(a,z,d)$$

3.6 Equilibrium Conditions

At time $t = T_{edu}$, given the initial distribution $H_{T_{edu}}(a,z,d)$, the equilibrium of the economy is characterized by a sequence of allocations $\{edu_t, o_t, c_t, a_{t+1}, k_t, l_t\}_{t=T_{edu}}^{T_{end}}$, factor prices $\{w_t, r_t\}_{t=T_{edu}}^{T_{end}}$, a tax rate $\{\tau_t\}_{t=T_{edu}}^{T_{end}}$ and the distributions of college and non-college graduates $H_t^c(a,z,d)_{t=T_{edu}}^{T_{end}}$ and $H_t^{nc}(a,z)_{t=T_{edu}}^{T_{end}}$ such that:

1. $\{edu_t, o_t, c_t, a_{t+1}, k_t, l_t\}_{t=T_{edu}}^{T_{end}}$ solves the individuals' policy functions for given factor prices.
2. Capital, goods and labor markets clear according to:

$$\int_{o_t(a,z)=entr} k_t dH_t^{nc}(a,z) + \int_{o_t(a,z,d)=entr} k_t dH_t^c(a,z,d) = \int a_t dH_t^{nc}(a,z) + \int a_t dH_t^c(a,z,d)$$

$$\int c_t dH_t^{nc}(a,z) + \int c_t dH_t^c(a,z,d) + \delta k_t = Y_t$$

$$\begin{aligned}\int_{o_t(a,z)=entr} l_t dH_t^{nc}(a,z) + \int_{o_t(a,z,d)=entr} l_t dH_t^c(a,z,d) = \\ \int_{o_t(a,z)=work} dH_t^{nc}(a,z) + \int_{o_t(a,z,d)=work} dH_t^c(a,z,d)\end{aligned}$$

with total output Y_t given by:

$$\int_{o_t(a,z)=entr} \left[e^{\xi_t^{nc}} (k_t^{\alpha} l_t^{1-\alpha})^{1-\nu} \right] dH_t^{nc}(a,z) + \int_{o_t(a,z,d)=entr} \left[e^{\xi_t^{c}} (k_t^{\alpha} l_t^{1-\alpha})^{1-\nu} \right] dH_t^c(a,z,d)$$

3. The budget constraint of the government clears, as outlined in Section 3.5.
4. The sum of incidental bequests (by non-surviving individuals) and bequests by the oldest generation alive in T_{end} covers the sum of the assets of the new generation of young adults:

$$\begin{aligned}&\int (1-\theta_t) a_t dH_t^{nc}(a,z) + \int (1-\theta_t) a_t dH_t^c(a,z,d) + \\ &b * \int (a_{T_{end}} dH_{T_{end}}^{nc}(a,z) + a_{T_{end}} dH_{T_{end}}^c(a,z)) = \int a_{T_{edu}} dH_{T_{edu}}(a,z,d)\end{aligned}$$

4 Quantitative Exercise

This section of the paper quantifies how much of the entrepreneurial differences across individuals with and without university education can be explained by the presence of student loans in the balance sheet of college graduates. I begin by estimating the model on the US economy using various sources of data, and then analyze the main quantitative predictions of my framework in terms of individual choices and aggregate outcomes. In the next sections, I will also investigate the impact of student loans bankruptcy availability on the extensive and intensive margins of entrepreneurship, and study how the rise in college tuition and premium has affected the increase in student debt and the decline in entrepreneurial rates of college graduates over the last decades.

4.1 Calibration

In what follows, I present the calibration strategy and discuss the quantitative fit of my framework with respect to targeted moments of the data. The reference period in the model is a year: agents that decide to attend college spend 4 years in university (i.e: $T_{edu} = 23$), then all individuals are active on the labor markets for 40 years more and retire at the age of 63 (i.e: $T_{work} = 63$), potentially living for 25 additional years in retirement (i.e: $T_{end} = 88$). Note that survival probabilities are set to reflect survival rates and life-expectancy for the US.[48] Of the 22 parameters I need to estimate, 9 are fixed outside of the model and summarized in Table 8. As standard, I use the coefficient of risk aversion $\gamma = 2$, the capital share $\alpha = 0.36$, and the yearly depreciation rate $\delta = 0.1$.[49] Secondly, I set the pension replacement rate in the model economy to be on average 50% of households' income in their last working period,[50] which is close to the one reported by De Nardi et al. (2020).

Table 8: Externally Fixed Parameters

Fixed	**Value**	**Description**
γ	2	Risk aversion
α	0.36	Share of capital in production
δ	0.10	Capital depreciation rate
p	0.50	Pension replacement rate
T_{repay}	15	Student loan repayment term
r^d	0.05	Interest rate on student loans
$\underline{d}$	\$9,800	Borrowing limit on student loans
s	\$5,625	College grant(s)
θ	(see text)	Survival probabilities

Furthermore, I set the length of the student loan repayment term to $T_{repay} = 15$, because, before 2010, almost the totality of college borrowers were enrolled in 10-years fixed repayment plans, which often extended to 20 years (see Abbott et al. (2019) and Daruich (2018) for a similar strategy). Plans that are instead tied to the income of individuals have been recently growing, but represent only 10-15% of the repayment programs subscribed in the last decade. I will nonetheless explore more in depth the difference between fixed and income-based repayment plans and their implication for workers and entrepreneurs in the next section. It is also important to mention that I calibrate several education-related parameters – r^d, $\underline{d}$, s – to reflect the current legislation on student loans and to follow common strategies already adopted in this literature, but their final values also depend on other internally fitted parameters, as it is explained in further detail below.

In particular, as in Abbott et al. (2019), I allow the scholarship term s to have two components, denoted by s_1 and s_2. The former is need-based and hence depends on individuals' financial resources, while the latter is proportional to students' merit. In particular, I assume the first grant to be inversely related to initial wealth and given by: $s_1 = \phi_1 a^{-\psi_1}$, while the second one to increase

[48] I directly take estimates from: `https://benjaminmoll.com`

[49] Commonly used values range from 0.06, as in Buera and Shin (2013), to 0.1, as in Clementi and Palazzo (2016).

[50] See `https://data.oecd.org/pension/net-pension-replacement-rates.htm`

with individuals' initial productivity according to: $s_2 = \phi_2 z$. Overall, the parameters ϕ_1 and ϕ_2 are calibrated so that scholarships cover respectively 15% and 10% of the average financial need of incoming students (shares are computed taking into consideration the mean amount awarded per student and the average share of students receiving it).[51,52] Furthermore, to replicate the progressivity of need-based programs – governed by the parameter ψ_1 in the model – I target the (negative) correlation between grant aid received and students' family income bracket.[53]

Table 9 reports all the other internally fitted parameters, which I proceed to discuss. First, I pick $\beta = 0.98$ to match an average annual interest rate for the US economy of $r = 4\%$.[54] I then set the wedge between r and the interest rate on student debt such that $r^d = 0.05$, in line with the average interest rate on education loans prevailing in the last decade.[55] The college tuition parameter χ is instead calibrated to replicate the share of the adult population with a college degree, which is considered to be around 35% over the last 10 years.[56,57] The estimated value $\chi = 1.25$ amounts to almost 30% of the median yearly income in the model (roughly \$25,000), consistent with recent US data on college costs and households' income. Accordingly, the lower bound on student loans $\underline{d}$ is set to over a third of the average yearly tuition for a 4-years college degree. Since I abstract from graduate studies, I compute an average of the maximum amount of education loans granted for undergraduate degrees across dependent and independent students, considering both subsidized and unsubsidized federal loans, which corresponds to almost \$10,000 per year.[58]

I also need to assign values to the parameters that define the initial distribution of wealth across the population and the correlation between initial assets and productivity upon birth. Assuming that the distribution of assets follows a log-normal shape, I normalize the mean to 1 and set the dispersion σ_a to match the fat right tail of the US wealth distribution, following recent estimates by Zucman (2019). Since the wealth agents are endowed with in period 1 influences college enrollment and the amount of student debt they choose, I check that the correlation between initial a and d in the model mimics the correlation between family contributions and student loans reported in Folch and Mazzone (2020).[59] Moreover, I calibrate the correlation between assets a and productivity z upon birth – denoted by ρ_{az} – to match the inter-generational persistence in earnings documented by Chetty et al. (2014) for the US economy.[60] Since my model does not feature households' dynasties, I instead compute the correlation between individuals' average (log) income over the life-cycle and their initial (log) assets, which can be interpreted as parental wealth.

Secondly, the span of control parameter is fitted such that the income share of the top 10% agents in the distribution of earnings is the same in the data and in the model. This choice is motivated by the fact that $1 - \nu$ regulates firms' scale of operations and, as a consequence, affects the profits of entrepreneurs, who are likely to belong to the top deciles of the earnings distribution. In that,

[51] See https://www.urban.org/urban-wire/what-better-data-reveal-about-pell-grants-and-college-prices.

[52] See https://www.usnews.com/education/best-colleges/paying-for-college.

[53] See the information reported for the 2015/2016 cohort at https://professionals.collegeboard.org/pdf/trends-spotlight-family-income-net-price.pdf. Abbott et al. (2019) adopt a similar strategy and target the progressivity in means-tested grants considering, however, data for the fiscal year 1999/2000.

[54] This figure reflects well the average interest rate prevailing in the US economy over the last 30 years.

[55] See https://educationdata.org/average-student-loan-interest-rate.

[56] See www.census.gov/newsroom/press-releases/2020/educational-attainment.html.

[57] The average tuition for 4 year degree is currently around \$112K, with the average debt at graduation being \$35K.

[58] See https://studentaid.gov/understand-aid/types/loans/subsidized-unsubsidized.

[59] Using US individual-level data from the Baccalaureate and Beyond Longitudinal Study, Folch and Mazzone (2020) report a correlation of 0.15 between family contributions and education loans. I compute a similarly moment in my model, and show that the correlation between initial wealth and the amount of student debt at graduation is 0.1391.

[60] A similar strategy is used in Daruich and Kozlowski (2020) to discipline inter-generational human capital.

Table 9: Internally Fitted Parameters

Fitted	Value	Description	Moment	Model	Data
β	0.98	Discount factor	Interest rate	0.04	0.04
χ	1.25	College tuition	Educational rate	0.37	0.35
σ_a	3.50	Dispersion initial wealth	Top10 wealth share	0.69	0.70
ρ_{az}	0.25	Correlation initial (a, z)	Inter-generational earnings	0.31	0.28
ν	0.78	Span of control	Top10 income share	0.45	0.45
σ_ϵ	0.305	St deviation prod shocks	Top25 employment share	0.63	0.65
ρ_z	0.92	Persistence entrep prod	Serial correlation revenues	0.84	0.80
λ	3.00	Financial constraint 1	Avg. corporate debt/GDP	0.30	0.35
η	0.15	Financial constraint 2	Δ Entr rates w/ – w/o Sloans	5pp	5pp
ζ_1^c	0.0573	Trend income growth (college)	Income growth year 0 - 30	0.84	0.86
ζ_2^c	0.0012	Curv. income growth (college)	Income growth year 30 - 40	0.07	0.05
ζ_1^{nc}	0.031	Trend income growth (no coll)	Income growth year 0 - 30	0.48	0.48
ζ_2^{nc}	0.0004	Curv. income growth (no coll)	Income growth year 30 - 40	0.08	0.10

I follow an extensive literature on income and wealth concentration in the US (see Batty et al. (2019) and Zucman (2019) for example), which shows that the top 10% richest Americans make up for almost 45% of aggregate earnings in the economy. My estimated value for the span of control parameter $1 - \nu = 0.78$ is close to the ones obtained by several other papers on US entrepreneurship.[61] As a robustness check, I can alternatively calibrate $1 - \nu$ to match the share of entrepreneurial wealth in aggregate wealth,[62] without changing the nature of my results.

To identify the volatility σ_ϵ of the entrepreneurial productivity shock, I target the employment share of the top 25% largest firms, computed using the 1980-2019 Business Dynamics Statistics dataset. A bigger σ_ϵ implies greater dispersion in the productivity process (by means of thicker tails in the distribution) and hence higher employment generation by large businesses.[63] My calibrated value $\sigma_\epsilon = 0.305$ is in line with the range of US estimates provided by Lee and Mukoyama (2015). In addition, I use a standard measure for the average serial correlation of revenues across US firms to identify the persistence ρ_z of the idiosyncratic entrepreneurial productivity process.[64]

Next, to calibrate the parameter λ, which governs the extent of firms' borrowing constraints, I match the average US non-financial corporate debt over GDP.[65] I focus on non-financial corporate debt because other measures of total (country's) debt merge together household and corporate liabilities, and hence cannot be mapped correctly into my theoretical framework.[66] In addition, I

[61] Values for the US typically range from 0.78 (see Buera and Shin (2013)) to 0.88 (see Cagetti and De Nardi (2006)).

[62] This is the calibration strategy followed by Cagetti and De Nardi (2006).

[63] Size is measured in terms of total employees, as also in Buera and Shin (2013) and Midrigan and Xu (2014).

[64] As discussed in Clementi and Palazzo (2016), estimates for ρ_z can be found to be as low as 0.8 and as high as 0.97. My final estimate $\rho_z = 0.92$ is similar to the one used by papers in this field such as Lee and Mukoyama (2015).

[65] See the entire series on FRED website: `https://fred.stlouisfed.org/graph/?g=VLW#0`.

[66] To pin down λ, I do not use SCF due to the lack of a proper variable capturing firm liabilities, which makes it difficult to compute debt-to-sales ratios. SCF only reports personal or family assets used as collateral for business purposes, and the mean ratio of collateralized assets to gross sales is between 0.60 and 0.70, with the median between 0.15 and 0.20. Due to this particular skewness, I would otherwise attribute too much weight to extremely high ratios. As an

use the relative percentage points (p.p) difference in entrepreneurial rates across college-graduate entrepreneurs with and without education loans to discipline the parameter η, which affects by how much outstanding student debt balances reduce the collateral that can be pledged by entrepreneurs on financial markets. (Unconditional) entrepreneurial rates for university-educated individuals with and without student loans are computed using SCF data for the last decade.

Finally, I have to calibrate 4 parameters related to the deterministic efficiency profile of agents with and without college over their life-cycle. Using SCF data, I set the values of ζ_1^c and ζ_1^{nc} to mimic the growth in the income profiles of US households with and without university degrees in the first 30 years of their working career. I then pin down ζ_2^c and ζ_2^{nc} targeting again the average growth in individuals' income profiles, but focusing instead on the last 10 years of their working life. The moments I compute for this final step of the calibration are close to those reported in Lagakos et al. (2018) across different sources of US data.[67] As shown in Table 9, the estimated values for ζ_1^c, ζ_1^{nc}, ζ_2^c and ζ_2^{nc} reflect the fact that income growth is faster at the beginning of the life-cycle of individuals, and instead slows down progressively as agents move towards retirement.

As a concluding remark to this subsection, it is relevant to mention that the tax rate τ pinned down in GE to balance government expenditures and revenues in the model economy is 0.2.[68] Moreover, the fraction of assets bequested by individuals upon their death, denoted by b, is such that the bequests of the last generation alive in T_{end} and the accidental bequests left by those who die before T_{end} cover the sum of the assets of newly-born cohorts.[69] This ensures that the new generations of young households in the model have all the same initial distribution of wealth.

4.2 Model Validation in the Cross-Section

In what follows, I discuss the quantitative fit of the model with respect to numerous dimensions of the SCF data that were not targeted during the calibration. First, I correctly predict that the share of borrowers among all college students is around 60%, as reported in recent US estimates. Secondly, the model replicates not only the general level of entrepreneurial activity in the economy, but also the relative composition of the entrepreneurial sample. It is important to stress that my estimation has only targeted the average share of educated individuals and the relative p.p. difference in entrepreneurial rates across college-educated entrepreneurs with and without student debt. The quantified model can instead match – as untargeted moments – the average US business ownership rate for the last decade,[70] as well as the average entrepreneurial rates of individuals with and without college. I also fit the fraction of entrepreneurs with and without student loans and without a

alternative, one can instead use Compustat, which covers publicly listed US firms between 1980 and 2016. The ratio of current liabilities to revenues is on average 0.41, in line with estimates from FRED data. Moreover, Morazzoni and Sy (2021) document a similar debt-to-sales ratio for a sample of US startups using the Kauffman Firm Survey.

[67] The values for ζ_1^c, ζ_1^{nc}, ζ_2^c and ζ_2^{nc} are consistent with the elasticities of salaries to age estimated by Daruich (2018), who also documents a steeper wage-profile for college vs non-college graduates using PSID data for the US.

[68] In the US, it is estimated that the average net income tax of single and married workers is 22% and 7% respectively (see for example `www.oecd.org/unitedstates/taxing-wages-united-states.pdf`).

[69] Since death may occur before T_{end}, the total sum of resources bequested in the economy corresponds to voluntary and accidental bequests. Voluntary bequests are precisely the calibrated share b of $a_{T_{end}}$ for those living till T_{end}, while accidental bequests are the assets accumulated by agents till the year $t < T_{end}$ in which they accidentally die.

[70] The average share of business owners in the 1989-2019 SCF sample is 0.13, down to 0.10 considering the last decade only. Self-employment rates are slightly higher for the same periods (0.14 and 0.12 respectively). Note that I am considering averages for individuals that are active in the labor force. Considering all survey respondents and hence computing population averages leads to an average entrepreneurial rate of 0.07. Similar statistics are reported by other sources, such as the OECD (see `https://data.oecd.org/entrepreneur/self-employed-with-employees.htm`).

college degree. For a comparison, Table 10 reports the moments computed in the model simulation and the ones documented empirically using the last 10 years of SCF data.

Table 10: Untargeted Moments

	Model	**Data**
Entrepreneurship & Education		
Share of Student Borrowers	0.55	0.60
Average Entrepreneurial Rate	0.12	0.10
Average Entrepreneurial Rate College	0.15	0.14
Average Entrepreneurial Rate Non-College	0.10	0.08
Share of Entrepreneurs with Student Debt	0.11	0.15
Share of Entrepreneurs without Student Debt	0.37	0.44
Share of Entrepreneurs without College	0.52	0.41
Model-Implied Elasticities		
Business Size (in Employees) to Student Loans	-0.641	-0.859
Business Profits to Student Loans	-0.210	-0.082
Business Sales to Student Loans	-0.169	-0.098

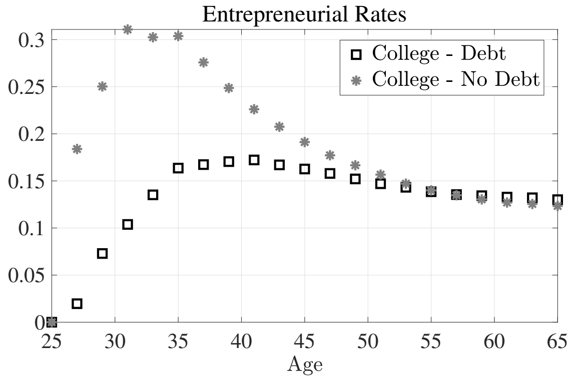
Considering households with a college degree and taking a life-cycle perspective, I can analyse several differences in entrepreneurial outcomes across graduates with and without education loans. In the data, having attended university predicts higher chances of becoming entrepreneurs: such association is typically attributed to higher human capital accumulation by college graduates, strong complementarities between education and labor market experience, and peer effects.[71] Similarly, in my model economy, having a university degree is positively related to undertaking entrepreneurship, due to the fact that college graduates face a higher deterministic efficiency profile throughout their life-cycle, regardless of their occupation. However, the repayment of student loans slow down the accumulation of wealth, while outstanding college debt balances lower the amount of collateral that can be pledged by entrepreneurs when renting capital on financial markets. The combination of these mechanisms can replicate the empirically estimated heterogeneities in the entrepreneurial rates of college graduates with and without education loans, as reported in Table 10. Before the debt is fully repaid, college borrowing discourages or delays entry into entrepreneurship, and, as shown in Figure 3, student borrowers see a catch up in business ownership rates between 15 and 20 years after completing college compared to graduates without loans.[72]

Keeping the focus on college-educated business owners, my model also predicts that firms of individuals with college loans are smaller than those of owners without student debt, and can secure less external funding. Considering the number of workers employed and the total sales or profits generated, I can match between 30 and 80% of the differences across entrepreneurs with

[71] See for example Michelacci and Schivardi (2020), Lerner and Malmendier (2013) and Van der Sluis et al. (2008).

[72] The deterministic growth in individuals' efficiency profile is also responsible for the growth in entrepreneurial rates over agents' life-cycle. In the data, the elasticity of business ownership rates to age is 0.0028 (netting out the effect of assets, demographic factors and year FE), while it is 0.0027 in the model economy. This result also highlights the importance of modeling entrepreneurial productivity as the combination of both a stochastic and deterministic component, with the latter precisely capturing the growth in skills and experience of households over their life time.

Figure 3: Extensive Margin



and without student debt. In particular, Table 10 collects the empirically estimated elasticities of business profits, sales and size to the amount of student debt owed by entrepreneurs.[73] Exploiting my calibrated model, I run equivalent regressions using a simulated panel of 50,000 households. A 1% increase in the amount of student debt owed decreases business profits and sales by 21% and 17% in the model, compared to the 8% and 10% elasticities computed empirically. Moreover, a 1% increase in the amount of student debt owed leads entrepreneurs in the model to hire on average 0.64 employees less, similarly to the 0.84 coefficient that has been estimated in SCF data.[74]

As a consequence of the tighter collateral frictions they face, I observe a stricter selection into entrepreneurship by college graduates with student loans. To exemplify this point, I compute the average product of capital (hereafter $arpk$) as the ratio between output and capital for entrepreneurs in the model. The $arpk$ is an indicator for how capital is allocated across productive units because, absent distortions, capital should flow similarly to entrepreneurs with and without student loans, ensuring no heterogeneity in $arpk$ across firms (see Hsieh and Klenow (2009)). Yet, student loans decrease the collateral that can be pledged on financial markets, which constitutes a barrier to the optimal allocation of capital across units that are more productive. Consistent with that, indebted college graduates that operate a firm have a 6% higher $arpk$, controlling for owners' assets. While I cannot compare such model-implied elasticity to any empirical counterpart due to the lack of data on capital in the SCF, this quantitative result implies that my calibrated economy features capital misallocation across firms run by entrepreneurs with and without student loans.

To further explore the extent of the entrepreneurial distortions generated by outstanding college debt, I analyse a counterfactual scenario in which I eliminate the difference in firm borrowing constraints across individuals with and without student loans. In practice, I set the parameter η to 0, recompute the equilibrium outcomes, and then compare the subsequent results to the baseline economy. First, whenever college debt does not contribute to tightening entrepreneurial financial frictions, the share of student borrowers increases by 1 p.p., and the average amount of debt taken for college scales up by 7%.[75] Second, indebted college graduates can leverage by more their own

[73]For the comparison with the model, I estimate again the regressions of Section 2 in the 2009-2019 SCF sample controlling for age, assets and education (and business size when the outcome variable is either sales or profits). I net out survey year FE to control for heterogeneous economic conditions across the years in the sample, as well as demographic characteristics that I do not explicitly model theoretically, such as gender, ethnicity and marital status.

[74]Model regressions on the simulated panel of households moderately overestimate the elasticities of business outcomes to student loans computed on SCF data for the last decade. A possible explanation for this result is that the model overpredicts by a third the average amount of college borrowing with which individuals graduate. Due to the non-linearities present in the model, this may imply relative higher barriers to entrepreneurship compared to the data.

[75]The college attainment rate goes up by only 0.3 p.p. in this counterfactual economy. This suggests that, if college

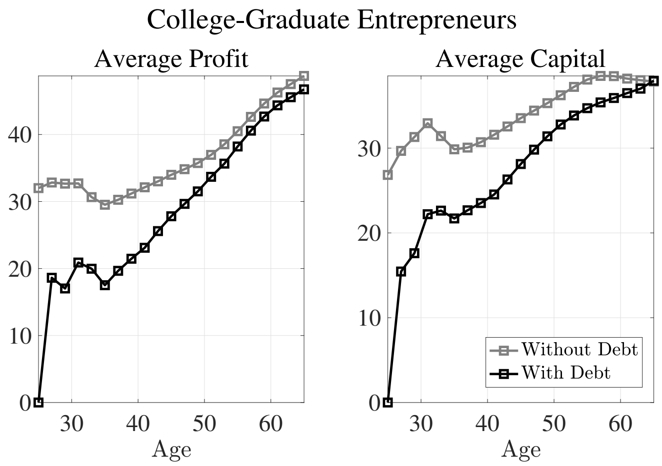
wealth in order to rent capital for business purposes whenever $\eta = 0$. As a result, their capital-to-labor ratio and entrepreneurial output increase by 4.96% and 5.39% respectively. Moreover, due to the improved allocative efficiency, aggregate production in the economy increases by 2.11%.

Table 11: No Difference in Entrepreneurial Constraints With and Without Student Loans

	Output (w/ Student Debt)	Business Debt (w/ Student Debt)	$\frac{\text{Capital}}{\text{Labor}}$ (w/ Student Debt)	Output Aggregate
Change wrt Baseline	+5.39%	+3.75%	+4.96%	+2.11%

As a final remark note that, in the model economy as in the data, entrepreneurial performance changes with individuals' age, due to assets accumulation and to the deterministic growth in agents' efficiency profiles. As a consequence of that, the gap in the average profit or capital between college-educated entrepreneurs with and without student debt decreases over time, especially after indebted households finish paying off their loans (i.e: 15 years after graduating college). However, since overcoming firm financial frictions through savings takes time, the gap in the average capital rented by college graduates with or without loans is wider and persists for relatively longer compared to other dimensions of business performance, as reported in Figure 4.

Figure 4: Intensive Margin



4.3 The Rise in Student Loans and the Decline in Entrepreneurship

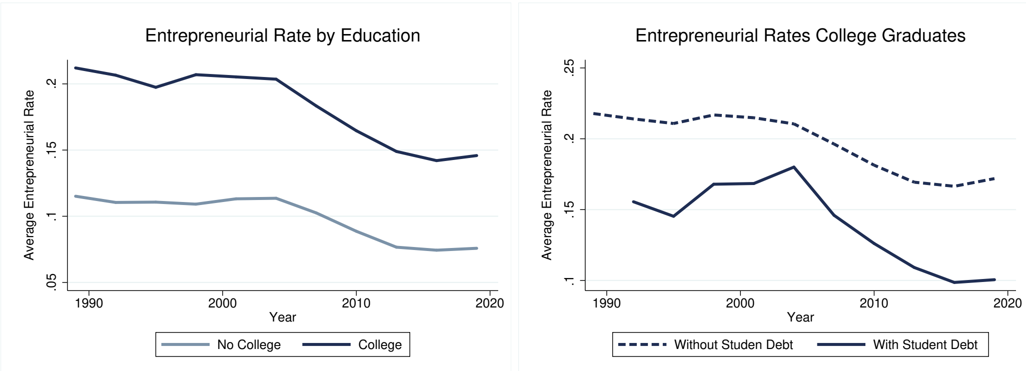
In what follows, I use the calibrated model to analyse the rise in student loans and the drop in the business ownership rate of college graduates over the last three decades. As documented by several previous contributions, US entrepreneurial rates and dynamism have steeply declined over the past 30 years (see Decker et al. (2014) for example). Using different household-level surveys, Jiang and Sohail (2017), Kozeniauskas (2018) and Salgado (2020) have also recently shown that the drop seems to have been bigger for college graduates, a phenomenon referred to as the "skill-biased entrepreneurial decline". Possible explanations for such a steeper decline in entrepreneurship for college graduates range from the advent of a skill-biased technological change to the steady fall

borrowing was not to influence entrepreneurial financial frictions, agents that choose to study would borrow more, but the increase at the borrowing margin would be relatively bigger than the increase in college enrollment itself.

in the price of capital, which may have decreased skilled entrepreneurship if one assumes capital and skilled labor to be complementary in production. Both channels could have increased the ratio of wages relative to entrepreneurial returns, pushing highly-skilled agents at the margin to select out of entrepreneurship. Parallel to that, rising entry costs and outsized productivity gains by large non-entrepreneurial firms could also be responsible for part of the decline in the business ownership rate of more educated people, as observed in Kozeniauskas (2018).

In addition to the previously discussed evidence and reasoning, Figure 5 further suggests that, among highly-educated individuals, the decline in entrepreneurship over the recent years has been even steeper for college graduates with student loans. Could then the rise in student debt and the fall in entrepreneurship – especially for skilled individuals – be related, and, if so, how could their co-movement be rationalized through the lens of the model developed in Section 3?

Figure 5: Entrepreneurial Rates Over Time



A prolific literature on entrepreneurial dynamism has already investigated various potential explanations for the drop in US firm ownership rates over time. It is beyond the scope of this section to offer an exhaustive review of previous works, and the goal of the model is not to account for all the proposed channels at once and disentangle their relative explanatory power. Based on the analysis done so far on the tighter constraints faced by firm owners with education loans, I rather aim to shed light on an alternative mechanism that could link the decline in the entrepreneurial rates of college graduates and the increase in student debt. To this end, I first observe that there are at least two important trends related to the growth in education loans over the past decades, namely the rise both in the college premium and in the tuition for college. As extensively documented in Goldin and Katz (2010), Heathcote et al. (2010) and Doepke and Gaetani (2020) for example, the gap in average salaries between college and non-college graduates has widened over time. Today, workers that hold a bachelor degree earn on average 20-25% more than in the late 80's relative to high school graduates.[76] Moreover, the average price to attend either public or private universities has more than doubled since the 1980s,[77] growing faster than US inflation.[78]

Connecting together these trends, increasing returns to higher education might have boosted the demand and hence the price for college. While recent research has shown the impact of other demand factors – such as expansions in federal student aid and rising parental transfers – on college prices (see Gordon and Hedlund (2020)), several papers have specifically related the rise

[76]Researchers point at the so called skill biased technological change as a possible reason for such a rise in the college wage premium. In my exercise, I nonetheless consider such change over time as exogenously given, as it goes beyond the scope of my project to investigate what caused the rise in the wage of skilled compared to unskilled workers.

[77]See `https://nces.ed.gov/fastfacts/display.asp?id=76`.

[78]See `https://www.bloomberg.com/news/articles/2021-10-25/college-tuition-cools-off-lagging-inflation-by-most-since-1970s`.

in university tuition to the increase in the college premium (see Jones and Yang (2016) and Fortin (2006)). Parallel to that, higher university prices have been held responsible for the soar in student debt (see Kim and Kim (2022)). In the exercise that follows, I link the expansion of student loans to the increase in the cost to attend university plausibly engineered by the rise of the college wage premium. Since, in my model, student loans affect disproportionately individuals that become entrepreneurs – especially due to the tighter financial constraints they imply – I explore what is the effect of carrying a higher student debt burden on different entrepreneurial margins.

I estimate the model to the US economy of the late 80's, using available data to inform the main parameters already summarized in Table 8 and Table 9 for the baseline economy. I then keep everything fixed and target the following adjustments: a (i) 20% growth in the college premium, induced by a 15% increase in the parameter governing the life-cycle profile of college graduates' efficiency ζ_1^c;[79] and a (ii) rise of 16 p.p. in the college attainment rate, which is accompanied by a 180% hike in the tuition for college captured by the parameter χ.[80] The former adjustment stimulates college enrollment through higher returns to education, while the latter ensures to deliver the average share of adults with a college degree in the US for the last decade, which is around 35%, as discussed in the strategy for the baseline calibration. I solve and simulate the economies of the late 80's and today: through the comparison of these two steady states, I can quantify the changes in entrepreneurial rates and outcomes attributable to the increase in college demand, college prices and student debt over time. Results for this counterfactual exercise are shown in Table 12.

Table 12: Changes between the late 1980s and Today

	Data	Model
Targeted		
College Premium	+ 20.0%	+ 20.0%
College Attainment	+ 16 p.p	+ 16 p.p.
Untargeted		
Total Student Debt	+ 788.90%	+ 689.50%
Share of Student Borrowers	+ 30.0 p.p.	+ 35.9 p.p.
Entrepreneurial Rate Overall	- 4.25 p.p.	- 0.50 p.p.
Entrepreneurial Rate College Graduates With Loans	- 5.47 p.p.	- 1.82 p.p.

Raising the college premium and the price to attend university leads to a consistent increase in the share of college graduates and, importantly, in the share of student borrowers. From the late 80's to today, the fraction of individuals taking up loans to finance their degree has gone up by roughly 30 p.p. in the data, compared to an increase of 35.9 p.p. in my counterfactual exercise. Computing the total amount of education loans in the two steady states, it is clear that the size of student debt in the economy has grown almost sevenfold, and this result lines up well with estimates from the

[79] Note that the increase in ζ_1^c brings about an increase in the life-cycle profile of efficiency for both entrepreneurs and workers: this is consistent with evidence from Michelacci and Schivardi (2020) showing that the college premium seems to have increased similarly for both entrepreneurs and workers in the US over the last decades.

[80] I consider changes in the average tuition from the late 80's till today, using constant 2019-2020 dollars, as reported by the US Department of Education at `https://nces.ed.gov/programs/digest/d20/tables/dt20_330.10.asp`.

Congressional Budget Office of the US Government.[81] The model can then match 1/10 of the decline in entrepreneurial rates for the overall population since the 1980s, and 1/3 of the drop in entrepreneurial rates for college graduates with loans over the same period.

As expected, the rise in student debt engineered by soaring college demand and prices can explain a rather small share of the overall decline in entrepreneurial dynamism, suggesting that many more forces are indeed at play in the data and can in fact rationalize the drop in firm ownership rates for both college and non-college graduates, as argued for example in Decker et al. (2014). Yet, the increase in education debt and in the share of student borrowers might have played a much more important role in determining the fall in entrepreneurial dynamism for indebted highly-educated individuals over the last decades. In this, my findings complement the results in Salgado (2020) and Jiang and Sohail (2017), who focus on the relationship between the growth in the skill premium and the skill-biased entrepreneurial decline. Specifically, in the presence of firm collateral constraints that depend on entrepreneurial pledgeable assets and are hence tightened by outstanding education loans, I highlight the role of higher college tuition and student debt in depressing business ownership rates, especially for indebted highly-skilled individuals.

5 Bankruptcy Availability

A cornerstone of US consumer credit markets are personal bankruptcy laws, which can provide loan discharge to distressed debtors under specific procedures.[82] Unlike other forms of consumer debt, student loans have become almost completely non-dischargeable in bankruptcy since 1998.[83] Exceptions regard individuals that join the public sector or the army, people affected by disabilities and debtors who can prove *undue hardships*. However, less than 0.001% of borrowers meet these standards and succeed in filing for bankruptcy (see Iuliano (2012)), while roughly 10% of outstanding student debt is currently in default.[84] As discussed by Yannelis (2016), policy makers are actively debating about re-allowing education loans discharge, and the White House has discussed reintroducing bankruptcy protections for student debt holders both in 2015 and 2018. From a macroeconomic point of view, the work by Ionescu (2011) has stressed the need to quantitatively study different bankruptcy regimes in the student loan market, and understand their implications for repayment incentives, human capital investment and aggregate welfare.

Why the availability of education loans bankruptcy could matter for entrepreneurship? Krishnan and Wang (2019) argue that student debt can reduce individuals' "tolerance for risk", including their propensity to open and run a business. By making bankruptcy unavailable, the 1998 reform could have increased the aversion of indebted college graduates to undertake entrepreneurial projects. But several other mechanisms may also apply. Before the reform took place, education loans were a type of unsecured debt that was easier to default upon, particularly if agents were facing financial hardships (see Yannelis (2016)). Student debt discharge might have then ensured households a "fresh start", especially because, in the US, credit risk scores are known to recover faster for bankrupt individuals compared to those remaining insolvent (see Albanesi and Nosal (2018)). According instead to the current legislation, borrowers who cannot repay or con-

[81]See `https://www.cbo.gov/publication/56754`.

[82]In 1934, US Supreme Court stated that bankruptcy "gives to the honest but unfortunate debtor a new opportunity in life and a clear field for future effort, unhampered by the pressure and discouragement of pre-existing debt".

[83]The *Higher Education Amendments* bill was first introduced in the House in January 1997, then it was approved by the House in May 1998 and by the Senate in July 1998, and it was finally put in place in October 1998.

[84]See `https://educationdata.org/student-loan-default-rate`.

solidate their education loans see their wages, income tax refunds or social security contributions garnished, and cannot abide to their student debt obligations or dismiss their outstanding balances.

Leveraging the fact that, before the 1998 Higher Education Act, student loans were dischargeable in bankruptcy after seven years in repayment, I analyse the impact of this reform on entrepreneurship in two steps. First, I establish a link between the 1998 bankruptcy reform and the outstanding student debt balances of individuals surveyed in the SCF up to a decade after. Since I have information on the repayment year in which they were at the time of the 1998 reform,[85] I am able to further distinguish respondents who had or had not access to education loans bankruptcy. Then, an RDD allows me to study the effect of outstanding student debt balances on entrepreneurship across cohorts who started repaying their education loans right before or after 1991.[86]

Secondly, I use the model from Section 3 to estimate the macroeconomic impact of the 1998 bankruptcy reform on individuals' entrepreneurial margins over the life-cycle, as well as on capital misallocation and aggregate US output. The goal is to replicate, in my calibrated framework, the key elasticity of business ownership to outstanding student debt that can be empirically estimated in the SCF data. Note that, in the model, outstanding loans affect firm outcomes mainly because of the borrowing constraint, which can be made less binding by the dismiss of student debt through bankruptcy. Then, analysing the partial equilibrium (PE) response of entrepreneurship to education loans discharge serves as a counterfactual and further validation of the quantitative fit of the model, particularly with respect to the parameter η, which captures the severity of the financial constraint imposed by outstanding student debt on college-educated entrepreneurs.

5.1 The 1998 Reform to Student Debt Bankruptcy

Before 1998, borrowers could discharge their student debt in bankruptcy after 7 years into repayment. The advent of the bankruptcy reform is not a pure randomized treatment, but yet a plausible source of exogenous variation in the repayment options and hence in the amount of student debt owed by affected individuals early in their career. The discontinuity in the availability of student loan bankruptcy by repayment year when the 1998 reform stroke can be exploited to estimate the impact of outstanding student debt on entrepreneurship through an RDD. As argued in previous paragraphs, there are several channels that could rationalize a potential effect of student debt bankruptcy on entrepreneurial margins. What is key to highlight is that the effect of loans discharge should first be reflected in a jump in the amount of student debt owed after the 7th repayment year for individuals who had the option to declare bankruptcy. Then, through the lens of the model introduced in Section 3, lowering outstanding student debt balances could subsequently impact entrepreneurial financial constraints and the choice to become an entrepreneur.

As a first step, I focus on the amount of outstanding student debt reported at the time of the survey interview by individuals that started repaying their loans at most 10 years before the 1998 bankruptcy reform was enacted. In the regression that follows, I control for the amount of debt agents graduated with and for the repayment year they were in by 1998 to account for the extent of initial loan balances and cohort effects. I also include as regressors households' demographics,

[85] It would be imprecise to instead focus on the graduation year of individuals, which may not coincide with the year in which loans start to be repaid due to grace periods and/or enrollment in post-graduate education.

[86] Through an OLS model, Krishnan and Wang (2019) find that individuals that graduated college after 1998 and took out student loans have a lower likelihood of becoming entrepreneurs. However, the bankruptcy reform also applied to graduates from previous cohorts, who graduated before 1998 but had not reached the 7th year of repayment. This motivates my different empirical strategy through an RDD based on the repayment year individuals were by 1998.

such as gender, ethnicity, marital and home-ownership status, as well as their income category by age and educational group as of the year in which they are interviewed. Focusing on the 10-years period after the introduction of the bankruptcy reform, Table B21 shows that agents that did not reach the 7th repayment year by 1998 are associated with higher outstanding student debt.

Having established some suggestive evidence on the relationship between bankruptcy availability and households' outstanding student debt balances, it is possible to investigate further the association between the 1998 reform and entrepreneurship itself. To begin with and as illustrated in Table B22, being past the 7th repayment year has a strong positive effect on the likelihood of becoming an entrepreneur for cohorts entering repayment before 1991, while it seems not to matter at all after 1991. It is important to stress that almost the totality of student debt repayment plans in the 1990s had a duration of 10 years. Accordingly, one should expect agents to have exercised the option to declare student debt bankruptcy right after reaching the 7th year into repayment, which is then the relevant cutoff to look at, as confirmed in the last 3 columns of Table B23. As such, I exploit the discontinuity in the availability of bankruptcy represented by the 7th year into repayment at the time of the 1998 reform to estimate the differential likelihood of becoming entrepreneurs for cohorts who started repaying before 1991, compared those who started at some given point between 1992 and 1997. I first run the following parametric probit regression:

$$Pr(BusOwner_{it} = 1) = F\left(\beta_0 + \beta_1 SubjectReform_i + \beta_2 \Delta_i^{cutoff} + \gamma' \mathbf{\Phi}_{it} + \alpha_t + \varepsilon_{it}\right) \quad (9)$$

where $BusOwner$ is a binary variable equal to 1 if individuals are entrepreneurs at the time of the survey, and to 0 if they are not. The regressor $SubjectReform_i$ is an indicator that takes a value of 1 if the respondent was before the 7th repayment year by 1998. Instead, Δ_i^{cutoff} captures how far from the 7th year cutoff individuals were by the time the reform stroke. Covariates and fixed effects are as in Appendix B. Results are shown below in Table 13 for different choices and combinations of bandwidths and control variables. Agents below the 7th repayment year cutoff by the time the reform was enacted are less likely to turn entrepreneurs later on in their life. This is true across different specifications: as reported in Table 13, widening the bandwidth around the 7th repayment year cutoff implies using more observations, which decreases the standard errors but can make the comparison across the treated and non-treated groups less accurate. The estimated β_1 coefficients range from -0.0568 to -0.0916, which corresponds to roughly a 1 p.p. decrease with respect to the average business ownership rates observed in the population.

The regression in Equation 9 is estimated parametrically, but coefficients do not qualitatively change when using the in-built Stata package from Calonico et al. (2015), which allows for more general specifications and data-driven choices of bandwidths.[87] In particular, Table 14 below shows similar results to Table 13. Individuals that were before the 7th repayment year of their student loans by the time the 1998 reform stroke are associated with a lower likelihood of becoming entrepreneurs.[88] Note that, across the different specifications, the algorithm optimally chooses regression bandwidths which span 4 time units above and below the 7th repayment year cutoff. In the baseline specification of Column (1), I use a linear polynomial to fit the regression. Esti-

[87] Such procedures optimize the bias-variance trade-off given the data in my sample.

[88] For roughly 700 individuals in the treated and control groups considered in Table 14 I have also information on the year in which they funded their business. Controlling for the initial amount of the student loan, the repayment year by 1998 and demographic characteristics such as gender, ethnicity, marital and home-ownership status and income category (by education and age group), the treated sample is associated with a 1.7 delay in their business funding year. Yet, the treated sample is also associated with higher sales and profits and bigger business size. This is consistent with selection into entrepreneurship becoming stricter across neighboring cohorts due to the removal of student debt bankruptcy provisions, which could have entailed lower net worth and a tightening of financial constraints.

Table 13: RDD Estimates of Likelihood of Business Ownership (Parametric)

	(2Yrs Bandwidth)	(2Yrs Bandwidth)	(3Yrs Bandwidth)	(4Yrs Bandwidth)	(4Yrs Bandwidth)
Subject to Reform	-0.0916***	-0.0901***	-0.0731**	-0.0772**	-0.0568**
	(0.0369)	(0.0369)	(0.0296)	(0.0261)	(0.0262)
Pre-Coll Controls	Y	Y	Y	Y	Y
General Controls	N	Y	Y	Y	N
Personal Wealth	N	Y	Y	Y	N
Survey Year FE	N	Y	Y	Y	N
Observations	1,565	1,565	2,168	2,887	2,887
R^2	0.0294	0.0472	0.0634	0.0487	0.0113
Avg Bus.Owners	0.1284	0.1284	0.1284	0.1284	0.1284

Notes: Estimates are average marginal effects. Robust standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1. Survey weights are used. The dependent variable is a binary indicator = 1 if the individual is a business owner, and = 0 if not. Pre-College controls refer to agent's gender and ethnicity. General control variables are agents' education, age, marital and home-ownership status, initial student debt amount and income or wealth category.

mates are stable when using a second order polynomial, when applying a uniform kernel function to weight the regressions (as opposed to the default triangular one),[89] when clustering standard errors at the repayment-year level and when introducing the same covariates as in Table 13.

Table 14: RDD Estimates of Likelihood of Business Ownership (Non-Parametric)

	Baseline	2nd Order Poly	Kernel(uni)	Clustered St.Errs	Covariates
Subject to Reform	-0.0632**	-0.0694**	-0.0691**	-0.0672***	-0.0657**
	(0.0316)	(0.0339)	(0.0305)	(0.0153)	(0.0313)
Observations	4,782	4,782	4,782	4,782	4,782

Notes: Robust standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1. Survey weights are used. The dependent variable is a binary indicator = 1 if the individual is a business owner, and = 0 if not. *Covariates* refer to agent's gender and ethnicity, age, marital and home-ownership status, assets and loan amount. Robust to include spousal income and the leverage ratio of the households instead of their asset positions.

As an additional robustness check, Figure B.1 illustrates graphically the discontinuity at the 7th year of student debt repayment and its relationship with the likelihood of business ownership later on in individuals' life. The relatively small sample size leads to modest jumps in the estimated coefficients at different repayment years, but Figure B.1 shows that the only significant discontinuity is represented by the 7th repayment year cutoff value. It is important to stress that the variable used to define the treatment and cutoff groups (e.g: the distance from the 7th repayment year by 1998) does not present jumps in its density around the relevant cutoff, suggesting little role for any confounding strategic behavior of individuals when approaching the 7th repayment year of their student loans.[90] Table B24 also illustrates the absence of any correlation between being in the treated group and the main covariates included in the estimation of Equation 9, which discards selection into the treated group. Finally, Table B25 contains and discuss standard placebos tests that strengthen and confirm the empirical validity of my RDD estimates.[91]

[89]The uniform kernel function gives the same weight to all observations.

[90]That would have been the case if households were to rush to open a firm right before the 7th repayment year of their student loans to be able to strategically discharge that debt as opposed to other loans, including business ones. The test run with the - rddensity - package in Stata returns a T-statistics of 7.5009, with a p-value of 0.0000.

[91]As an additional check, I run similar RDD specifications to investigate whether the availability of student debt

5.2 Macroeconomic Impact

The next step is to evaluate the impact of bankruptcy availability on entrepreneurship in my model economy. As in Kaboski and Townsend (2011), Lagakos et al. (2018) and Buera et al. (2021a), this counterfactual exercise is carried out in a PE setting, namely without recomputing aggregate prices. To remain close to the spirit of the RDD specifications presented above, my goal is to estimate the elasticities of the extensive and intensive entrepreneurial margins to the provision of a bankruptcy scheme, without letting the surrounding economic environment change at the same time. Nevertheless, it is also plausible to assume that the general equilibrium effects of the 1998 reform might have not been sizable. This is due to the fact that bankrupt individuals were a low share of all student debt borrowers back in the 1990's, and college graduates used to carry smaller balances than today: in particular, the average size of education loans upon graduation was \$10K for the 1992-1993 cohort, which is less than a third of what is computed for recent cohorts.

I hence fix the model parameters to the estimates discussed in Table 8 and Table 9, keeping input prices and the tax rate to their baseline values. Secondly, I simulate an alternative economy where I allow for student debt discharge after 7 years into repayment – the same way the option of bankruptcy on education loans was implemented before the 1998 reform. Moreover, when households in the model liquidate their student debt, they are required to use any available assets to cover for as much as possible the amount of defaulted college borrowing. This means that bankrupt individuals have to pay a sum equal to $\max\{a_t - d_t, 0\}$ in the year $t \geq 7$ in which they discharge their student debt. After that, they become free of obligations on all remaining balances and are no longer responsible for loan repayments. Note that, to get correctly the overall impact, I have to carefully replicate in the model economy the average share of student debt that used to be discharged before the 1998 reform. As mentioned in the previous paragraph, roughly 1-1.5% of education loans were dismissed in bankruptcy per cohort before 1998 (see Yannelis (2016)).

Two clear effects of student debt bankruptcy are worth discussing: on the one hand, college graduates who discharge their education loans after 7 years into repayment are then able to accumulate higher assets, as they become free of repayment obligations. At the same time, under the assumption that bankruptcy comes at no extra cost,[92] entrepreneurs' borrowing constraint may become less tight, leading them to rent higher levels of business capital. Both mechanisms are expected to boost the entrepreneurial rate of households with a college degree and who took out education loans to finance it, and to increase the amount of capital they rent for their business. The impact of bankruptcy availability on different entrepreneurial margins is reported in Table 15.

Bankruptcy availability boosts the entrepreneurial rate of college graduates with loans. Exploiting the 1998 reform, RDD regressions have found an elasticity of business ownership to student debt discharge between 6 and 9%. The model delivers a 7.64% coefficient, fitting more than 70% of the empirical estimates. The effect in the simulated economy is larger than in the data, consistent with the fact that bankruptcy availability is assigned randomly in my counterfactual, and I do not allow selection into student debt discharge along relevant individuals' characteristics. Moreover, bankruptcy availability increases business funding for college-educated entrepreneurs with student loans by 16.97%. By loosening their collateral constraints and expanding their capital rental

bankruptcy potentially affects other individuals' outcomes, such as their likelihood of marrying and/or being homeowners. Consistent with the discussions in Folch and Mazzone (2020) and Ji (2021), I overall find that outstanding student loans reduce the likelihood of buying a house and marry. All results are available upon request.

[92] This is not a straightforward assumption to make, as individuals declaring bankruptcy in the US are typically assigned a bankruptcy flag by banks, which lasts on their records for maximum 10 years. However, as found by Cohen-Cole et al. (2013), more than 90% of bankrupt individuals tend to receive credit shortly after filing for bankruptcy.

Table 15: Effect of Bankruptcy Availability on Entrepreneurship

Outcome	Change wrt to Baseline
Entrepreneurial Rate of college graduates w/ Student Loans	+ 7.64%
Data	*[6.32 – 9.16%]*
Entrepreneurial Debt of college graduates w/ Student Loans	+ 16.97%
Total Entrepreneurial Output	+ 0.60%

capacity, the option of discharging outstanding student loans after 7 years into repayment reduces capital misallocation in the economy. Finally, the counterfactual exercise shows an increase in aggregate entrepreneurial output and welfare of +0.60% and +0.05% respectively.[93]

6 Policy Counterfactuals

In what follows, I analyse four policy experiments to further investigate the interplay between the sources and characteristics of college financial aid and the extensive and intensive margins of entrepreneurship. I first study the effect of increasing the provision of college grants – considering need and merit-based ones – which represents a shift in the composition of total aid awarded to students. Second, I raise the borrowing limit on education loans. In the third experiment, I instead compare the baseline economy – where student debt repayment plans are *fixed* – to a counterfactual scenario where they become *income-driven*. In this case, indebted graduates pay the minimum between the fixed amount and a given share of their income. This different scheme makes payments less binding in bad times, and also implies that remaining student debt balances are forgiven and covered by public expenditure after their repayment term expires. Note that all interventions are supported through potentially higher taxation by the government, whose budget constraint must balance in every period. Moreover, by fostering college enrollment, all three policy experiments could in principle increase the amount of student debt per borrower and the size of the market for education loans, which makes their aggregate effects a priori unclear.

Finally, I conduct a preliminary assessment of US President Biden's proposal to cancel part of outstanding student loans. Following the plan recently outlined by the White House, I simulate the introduction of debt relief in my baseline economy, and compute the impact that this one-time intervention could have on entrepreneurial margins and on the fiscal burden of affected taxpayers.

[93] The amount of student debt discharged in bankruptcy should in principle become a financial burden for the government in my framework, who would have to increase taxation to cover bankrupt student loans and meet its budget constraint. In turns, higher taxes would decrease consumption and savings, while higher entrepreneurial rates would raise the demand and price of capital and labor. The resulting impact on output and welfare could be ambiguous but can be quantitatively estimated allowing for the GE response. Without endogenizing the decision to declare bankruptcy, exogenous student debt discharge would increase the tax rate by 0.5 p.p., while the entrepreneurial rate of graduates and the amount of business debt they can secure would increase by less compared to the PE case (+7.63% and +15.28% with respect to the baseline economy). Note that college-graduate entrepreneurs have a higher efficiency profile and marginally crowd out non-college graduates from the entrepreneurial pool. Yet, the increase in entrepreneurial entry for college graduates would boost the demand for both capital and labor, raising equilibrium prices and depressing production. Combined with the increase in the tax rate, this would still decrease entrepreneurial output by 0.5%.

6.1 Expansion of Grants and Borrowing Limits

As a first remark to these policy counterfactuals, recall that my baseline model is characterized by missing markets for insuring against being born from a poor family (eg: having low initial wealth a). Moreover, neither the maximum amount of student debt individuals can borrow, nor the grants provided by the government can fully cover the tuition for college. As noted in Abbott et al. (2019), government interventions that guarantee an easier access to student loans or increase university subsidies can partially address such under-investment in higher education. In particular, they can ensure that a larger share of the population, especially highly-productive constrained individuals, benefits from higher income growth over their life-cycle. In turn, their higher income tax contributions can in principle compensate for the resulting increase in public expenditure.

I begin by analysing the effects of a potential expansion of grants schemes, noting that the scholarship term s in the baseline economy has two components, given by s_1 and s_2. The former is means-tested and depends on one's family available resources, while the latter is proportional to students' merit. In particular, I assumed the first grant to be inversely related to initial wealth and given by: $s_1 = \phi_1 a^{-\psi_1}$, while the second one to increase with individuals' initial productivity according to: $s_2 = \phi_2 z$. Together, these two scholarships account on average for 25% of the yearly cost of attending college. In the next counterfactuals, I follow the spirit of the policy exercise carried out in Abbott et al. (2019) and double the share of tuition covered by grants (from 20 to 40%).[94] Then, I examine the subsequent changes on education choices, entrepreneurial margins, aggregate output and welfare.[95] Both exercises are performed in GE, by recomputing r, w and τ.

Table 16: Expansion of Means-Tested Grants

	Entrepreneurship (w/ Stud. Debt)	Output (w/ Stud. Debt)	Business Debt (w/ Stud. Debt)	Output Aggregate	College Attainment	Welfare Aggregate
Change wrt Baseline	+6.00%	-0.8%	-1.24%	+0.14%	+6.02%	0.72%

First, I expand need-based scholarships by raising the parameter ϕ_1 from 2.20 to 2.25 and keeping instead fixed the progressivity of the subsidy with respect to students' wealth, captured by ψ_1. In so doing, the share of college tuition covered by need-based grants goes up from 12.5% to 25%. Second, I analyse an increment of merit-based scholarships by increasing the parameter ϕ_2 from 0.04 to 0.06, which raises the fraction of college tuition covered by merit-based grants from 10% to 15%. In both cases, the government has to meet the rise in public expenditure with larger fiscal revenues. However, the model economy is characterized by a higher income profile for college graduates and by a proportional income tax. Therefore, the increase in the share of college-educated individuals induced by a more generous provision of educational grants enlarges the amount of fiscal revenues collected by the government as well. Both policy changes analysed here are in fact fiscally self-sustained and do not lead to a higher equilibrium tax rate with respect to the baseline economy. Results from these counterfactuals are shown in Table 16 and Table 17.

Expanding need and merit-based aid raises by 6% and 10% the college attainment rate in the counterfactual economies, and by 19% and 25% the average amount of student debt per person. Increasing merit-based grants induces relatively productive but constrained students to substitute

[94] Also, note that my framework does not include any disutility or psychic costs of attending university. An excessive increase in the provision of grants could produce a counterfactual and unrealistic rise in college enrollment.

[95] I define aggregate welfare as the sum of utilities over consumption across the distribution of all individuals.

Table 17: Expansion of Merit-Based Grants

	Entrepreneurship (w/ Stud. Debt)	Output (w/ Stud. Debt)	Business Debt (w/ Stud. Debt)	Output Aggregate	College Attainment	Welfare Aggregate
Change wrt Baseline	+8.97%	+12.29%	+0.86%	+2.30%	+9.84%	+1.82%

education loans with publicly-provided financial aid (note that the share of student borrowers decreases by 4.9%), and allows a larger fraction of young adults to secure a high efficiency life-cycle profile through college education. Both mechanisms raise entry into entrepreneurship and the output of indebted college-educated business owners. Due to a higher share of college graduates within the entrepreneurial sample, the productivity cutoff to open a firm shifts rightwards: together with the aforementioned effects, the crowding out of marginally less productive owners increases aggregate output by 2.3%, and contributes to a 1.8% increment in aggregate welfare.

On the contrary, doubling the size of need-based grants does not equally succeed in attracting potentially constrained but productive students into college, and has in fact the downside effect of marginally increasing by 5.3% the share of indebted graduates. Overall, the larger fraction of borrowers and the higher student debt burden worsen – instead of improving – the entrepreneurial performance of college-educated agents with loans. As a consequence, the counterfactual economy under higher means-tested grants does not register substantial positive compositional effects within the entrepreneurial sample, which limits the gains in aggregate output and welfare.

In a third exercise, I examine the effect of loosening college borrowing limits and allowing students to take out larger loans to finance their degree. In the baseline economy, the lower bound on student loans $\underline{d}$ was set to a third of the average yearly tuition for a 4-years college degree. Since the model abstracts from enrollment in graduate studies, I considered an average of the maximum amount of loans granted for undergraduate degrees across dependent and independent students, including subsidized and non-subsidized federal loans, which corresponds to roughly \$10,000 per year. In the following experiment, I increase by 25% the maximum amount students can borrow to finance college, and assess the impact of this reform on education choices, entrepreneurial margins, aggregate output and welfare. Differently from the counterfactual in Abbott et al. (2019), I do not assume student loans to fully cover the tuition, but rather look at a middle-ground case.

Table 18: Expansion of Borrowing Limits

	Entrepreneurship (w/ Stud. Debt)	Output (w/ Stud. Debt)	Business Debt (w/ Stud. Debt)	Output Aggregate	College Attainment	Welfare Aggregate
Change wrt Baseline	- 9.38%	-3.09%	-4.98%	+0.82%	+16.67%	+1.89%

As shown in Table 18, the expansion of the borrowing limit on student debt results in a 17% rise in the educational attainment of young adults, and in a 58% increase in the average amount of borrowing per student. However, the share of borrowers among college graduates does barely move: this indicates that some prospective students are credit constrained, but the maximum amount of education loans granted does not affect one to one their discrete choice of taking out loans. Moreover, in line with the fact that student loans affects entrepreneurial margins in my model, a higher debt burden decreases business ownership for agents with student loans, limits the amount of business credit they get to finance capital acquisition, and reduces their entrepreneurial output. As a consequence of that, allocative efficiency – measured by college-educated entrepreneurs' $arpk$ – worsens in the economy by 6%. Yet, the economy now features a larger share of highly-educated entrepreneurs, who enjoy a higher efficiency profile and income growth over the life-cycle. Despite a 0.3 p.p. increase in the GE tax rate to support the expansion of student debt limits, and the negative effects registered in the early stage of indebted graduates' entrepreneurial careers, this policy change still results in almost 1 p.p. higher aggregate welfare and output.[96]

As a final remark to these exercises, I want to stress two key limitations of my framework: on the one hand, I have modelled the college premium as an exogenous factor. Specifically, I have assumed that university-educated individuals enjoy a given higher efficiency profile over the life-cycle, which is not affected by the increased supply of college graduates induced by the policy changes explored above. On the other hand, the price to get a degree – denoted by χ in the model – does not move either throughout these counterfactuals, despite the fact that expanding the provision of grants or the borrowing limit on student debt causes a rise in college demand.

In future work, a possible solution to tackle these issues would be to endogenize the supply-side of higher education or to include a college market, as in Cai and Heathcote (2022). At the same time, I could also microfound further the presence and extent of the college premium by assuming skilled and unskilled labor to have a different degree of complementarity to the technology in the production function of entrepreneurs, as in Salgado (2020). This would in principle allow the premium and the price for college to react to changes in the supply of highly-educated agents and in the demand for higher education itself, which in turn could affect the quantitative results of the policy experiments I have studied. For now, I follow Abbott et al. (2019), and keep the price and the higher efficiency profile induced by college education fixed when assessing the impact of reforms to university financial aid, observing that the share of college graduates does not in fact increase exponentially or unrealistically in the counterfactual policy scenarios that I analyse.

6.2 Income-Based Repayment Plan

The second type of policy exercises I perform is a change to the repayment structure of student loans, by making their repayment tied to the income of borrowers in any given year. In the US, there are currently four different types of income-driven plans, which include the Revised Pay As You Earn (REPAYE), the Pay As You Earn (PAYE), the Income-Based Repayment (IBR) and the Income-Contingent Repayment (ICR).[97] All of them entail a repayment that varies between 10% and 20% of agents' discretionary income.[98] Moreover, if the original loan is not paid off entirely after 20 or 25 years, depending on the plan, outstanding balances are forgiven. Interestingly, despite the fact that the US administration has passed actions requiring matriculating students to be informed about income-driven repayment options, these represent less than 15% of the plans subscribed in the last years. Similarly to Luo and Mongey (2019), I introduce in my model an IBR program, which was first launched in 2009, and assess its effects on macroeconomic outcomes.

Recall that, in my baseline economy, the initial loan balance due in repayment is divided into fixed tranches, which individuals pay along with interest rates on top of their outstanding debt

[96] The gain in individuals' efficiency induced by attending university is such that college-educated entrepreneurs, especially those without student debt, are relatively more productive than in the baseline economy, which raises both output and capital and labor demand. Higher input prices induce churning within the entrepreneurial sample, and make it harder for non-college individuals and college-graduates with student debt to open and run a firm.

[97] See https://studentaid.gov/manage-loans/repayment/plans/income-driven.

[98] See https://www.census.gov/library/publications/2021/demo/p60-273.html.

until the end of their repayment term $T_{repayFIX} = 15$. Next, I assume instead that student loans get repaid through an IBR plan, under which agents have to disburse the minimum between the fixed repayment amount and 15% of their current income, as long as it exceeds 150% of the poverty line established by the government. If the latter condition is not met, the repayment due is zero. Moreover, borrowers have to pay interests on their outstanding loan balances, as for the standard repayment plan, provided that these do not exceed the amount of the principal payment. For now, I do not allow either for the endogenous choice of repayment plan upon graduation, or for the option to switch between plans. As for the previous exercises, this counterfactual is carried out in GE, and the government covers with fiscal revenues any higher public expenditure caused by unpaid interests and debt forgiveness after 25 years into repayment (recall that $T_{repayIBR} = 25$).

As noted by Luo and Mongey (2019), individuals that carry low amounts of student debt may benefit from the standard repayment plan, which enables them to run down quickly their small balances without bearing the burden of large interest rate payments that is involved in longer IBR plans. On the contrary, the IBR program is preferred at moderately higher debt levels, as it ensures higher consumption early on in agents' careers, when income is lower and the marginal utility of consumption is higher. The effect on aggregate outcomes is hence hard to assess a priori, as it depends on the endogenous selection of individuals into education and student debt, and is potentially interlinked to their consumption-saving and occupational decisions over the life-cycle.

Table 19: Income-Based Repayment Plan

	Entrepreneurship (w/ Stud. Debt)	Output (w/ Stud. Debt)	Business Debt (w/ Stud. Debt)	Output Aggregate	College Attainment	Welfare Aggregate
Change wrt Baseline	-12.32%	+11.60%	+9.97%	+2.65%	+8.33%	19.29%

Table 19 shows that, if all prospective students were enrolled in IBR plans, the college attainment rate would increase by more than 8%, the share of borrowers would raise by 35 p.p. and the average amount of student debt per person would double. Since the government would have to cover unpaid interest rates and guarantee debt forgiveness after 25 years of repayment, the average GE tax rate would increase by 1 p.p. for all agents.[99] Results from this counterfactual exercise show that switching completely to IBR plans might not foster entrepreneurial entry, but may reduce the gaps in entrepreneurial outcomes across college graduates with and without loans. This is due to strong income effects, stemming from the fact that a longer repayment period and larger payments towards the middle-end of individuals' working careers reduce wealth accumulation, and discourage undertaking risky entrepreneurial activities. However, since student debt payments can effectively be delayed in bad times without increasing outstanding balances, adopting an IBR allows indebted college-educated entrepreneurs to rent higher capital and produce more on average.[100] The increase in business earnings and wages for all individuals in the economy more than compensate the higher fiscal pressure, and result in a 19% increase in welfare overall.

[99] As noted in Abbott et al. (2019), policy reforms may have upfront costs for longer term benefits to future generations. In particular, this could imply that some of these policies would be better financed using long term government debt, instead of taxes falling on current generations. I leave the consideration of transitional effects for future work.

[100] My results are in line with the analysis of Ionescu (2009), who shows that student debt repayment flexibility increases enrollment, decreases default rates, and lead to redistributional effects that benefit low-income households.

6.3 Student Debt Relief

In the last exercise of this section, I conduct a preliminary assessment of President Biden's recent plan to cancel off part of outstanding student loans, which was formalized and released by the White House on August 24$^{\text{th}}$, 2022. While the proposal for a potential student debt relief has been discussed at least since the last US presidential campaign, the debate recently re-gained momentum, as America's working families are starting to recover from the strains associated with the COVID-19 pandemic. Behind this intervention lies the belief that the cost of college borrowing has become a burden preventing most student debt holders from enjoying the advantages post-high school education should grant. In particular, Biden's Administration has stressed how middle-class borrowers struggle with high monthly payments and ballooning balances, which make it harder – in their words – to build wealth, buy a house, open a business or save for retirement.[101]

President Biden's proposal entails up to \$20,000 in debt cancellation for Pell Grant recipients with loans held by the Department of Education, and up to \$10,000 in debt cancellation for non-Pell Grant recipients, provided that they do not belong to the top 5% income earners. Since it is estimated that nearly every Pell Grant recipient came from a family that made less than \$60,000 a year,[102] student debt cancellation should specifically target low and middle-income debt holders, and would provide relief to 43 million individuals, including forgiving the remaining balance for roughly 40% of the borrowers. Yet, President Biden's forgiveness plan will lead the government to cover the \$400 billions cost through tax increases, spending cuts, borrowing or a combination of these tools. This has spurred a debate over the redistributional consequences of the proposed intervention, and part of the public opinion argues that this measure does not address the root cause of why students graduate from college with such huge debt burdens in the first place.[103]

In the counterfactual that follows, I simulate the introduction of President Biden's student debt relief in my baseline economy. Rather than offering a comprehensive discussion over whether and how to optimally conduct education loans cancellation, this exercise aims to study the impact that this particular intervention would have on specific aggregate outcomes. In practice, I take the calibrated steady state of my model and shock every cohort alive with a one-time student debt relief that mimics President Biden's proposal, and that affects individuals who are repaying their college borrowing. The set up of my framework allows me to closely replicate President Biden's plan, as Pell Grants are proxied by the need-based scholarship s_1, income is computed from either labor or entrepreneurial earnings, and a \$10,000 debt cancellation corresponds to roughly a third of the average education loans balances at graduation. Since, at the moment of writing, there is still uncertainty over the steps that will follow this initial relief, I describe and focus only on the choices and outcomes of affected cohorts. In particular, I analyse how the extensive and intensive margins of entrepreneurship may react, and what the resulting fiscal pressure from this measure would be, without recomputing market clearing conditions. Results are shown in Table 20.

Introducing a one-time student debt relief in the steady state of my economy would wipe out outstanding college loans for 52% of the borrowers (compared to the White House projection of 40% previously mentioned). In the cross-section, I also observe an increase of 2.11% in the entrepreneurial rate of college graduates, and a rise of 0.78% in the capital-over-labor ratio of firms

[101] See https://www.whitehouse.gov/briefing-room/statements-releases/2022/08/24/fact-sheet-president-biden-announces-student-loan-relief-for-borrowers-who-need-it-most/.

[102] See studentaid.gov/data-center/student/title-iv.

[103] See for example www.forbes.com/advisor/personal-finance/who-pays-for-student-loan-forgiveness/.

Table 20: Student Debt Relief

	Entrepreneurship (w/ Stud. Debt)	Capital-to-Labor (w/ Stud. Debt)	Interest Rate (Estimated Increase)	Avg Tax Rate (Estimated Increase)
Change wrt Baseline	+2.11%	+0.78%	+4.23%	+0.66 p.p.

run by college-educated entrepreneurs. Both effects may in principle represent improvements along the extensive and the intensive margins of entrepreneurship for indebted college graduates.

It is important to stress that these figures constitute an upper bound to the potential entrepreneurial gains that the current student debt relief proposal could obtain. Specifically, my preliminary assessment does not consider any feedback effect generated by subsequent changes in GE prices. Lower college debt repayments would free up more resources for all individuals and raise capital supply, but a higher entrepreneurial participation by graduates with loans would boost the demand for capital. On average, this could in fact increase the equilibrium interest rate by roughly 4%. Moreover, student loans forgiveness has to be covered exclusively with taxes, since the government in my model cannot issue public debt and uses only fiscal revenues to balance its budget constraint. As such, I estimate that college debt relief could increase the average fiscal pressure for all agents by almost 1 p.p., but nonetheless leave for future research a more thorough analysis of the full equilibrium response of the economy to President Biden's proposal.

7 Conclusion

In this paper, I have investigated the interplay of education and occupational choices over the life-cycle of households, focusing on the effect of student debt on entrepreneurship. Using micro-level data from the US Survey of Consumer Finances for the 1989-2019 period, I have documented a negative relationship between student loans and entrepreneurial outcomes. Specifically, individuals carrying student debt balances and who took out loans to finance their college degree are less likely to become business owners and obtain external funding. Their firms also tend to be smaller in size and to generate less revenues and profits, but they do present better profitability margins.

I have rationalized my findings into a GE heterogeneous agents model, where individuals differ by wealth, productivity, age, education and student debt. During youth, households decide whether to attend college and how much to take out in education loans. College gives them a income premium through higher (deterministic) efficiency growth, and student debt has to be repaid upon graduation. During their adult life, all individuals make occupational choices and decide whether to open a firm or be workers. When in repayment, education loans slow down the accumulation of wealth of college graduates, and tighten the borrowing constraint of indebted entrepreneurs. Calibrated to the US, my model replicates between 30 and 80% of the empirical differences across entrepreneurs with and without education, and with or without student debt.

Secondly, I used the 1998 reform to education loans bankruptcy to establish a causal link between student debt and entrepreneurship. Using SCF data and an RDD, I have estimated an elasticity of business ownership rates to education loans between 6 and 9%. I have then expanded my quantitative framework to include and allow for student loans bankruptcy under the legal terms in order before 1998. I found a 7.64% PE elasticity of entrepreneurship to student debt bankruptcy, which replicates closely its empirical counterpart. In such counterfactual scenario, capital misallocation would decrease, and entrepreneurial credit and output in the US would increase.

Finally, I have also used to model to investigate the relationship between the rise in college costs and student debt and the decline in entrepreneurship over the past 40 years. Specifically, I have shown that the boom in college demand and prices engineered by the rise in the college wage premium can account for the exponential increase in the amount of student debt per person and in the share of borrowers. In turns, higher student debt levels are responsible for 1/10 of the overall decline in entrepreneurship in the US, and for 1/3 of the decrease in business ownership rates for college graduates with loans. The model has also served as a quantitative laboratory to assess the effect of specific public policies on individuals' choices and aggregate outcomes. In particular, I have studied the impact of college aid expansions and income-based student debt repayment plans on entrepreneurship, capital allocation and aggregate productivity in the US. In future works, I believe it would be important to endogenize the supply-side of education and the college premium in the model. This would allow to analyse the equilibrium response of university demand and prices to shocks affecting the technology of the firms, or investigate how changes in the educational system manage to propagate and influence individuals' labor market outcomes.

Appendix

A Data Appendix

A.1 Variable Definition and Datasets Comparison

Table A1: Description of Demographic Controls

Variable	*Description*
Age	Age of the household (25 to 65 years old).
Ethnicity	Ethnicity of the household (White, Black, Latino, Other).
Education	It is a categorical variable measuring the highest level of education attained by owners. The original scale is from 1 (less than 4th grade) to 12 (professional school or doctorate). When specified, they are recoded into two levels, namely high school (and lower) and college (and higher) level. The latter refers to education categories "some college, but no degree", "associate's degree" and "bachelor's degree", "master's degree" and "professional school or doctorate".
Marital status	It is a binary variable equal to 1 if the household is married.
Number of Kids	Total number of kids in the household (0 to 10+).
Personal Debt	Includes principal residence debt (mortgages and HELOCs), other lines of credit, debt for other residential property, credit card debt, installment loans,and other debt.
Personal Assets	The sum of financial assets and non-financial assets held by households, such as savings account, bonds, annuities, retirement accounts, residences, vehicles among others.
Spouse Income	Income of working spouse, either from employment of self-employment
Home-Ownership	It is a categorical variable equal to 1 if households own the house where they live, and to 0 otherwise.
Parents' Education	It is a categorical variable measuring the educational attainment of the father and the mother. The levels are "less than high-school", "high-school diploma", and "college degree".

In Table A1, I describe the variables used in the main regressions of the paper, which refer to individuals' demographic characteristics, their average income or financial position. Note that Table A2 and Table A3 define instead the variables related to the businesses run by respondents and to their student loans. In Table A4, I also offer a comparison between the SCF and other datasets used to investigate trends and patterns in US college borrowing over time and within student cohorts. In particular, aggregate statistics from SCF related to the share of borrowers within college recipients and to the civilian population are compared to those obtained with data from (i) the National Center of Education Statistics (NCES), which includes surveys such as the National Postsecondary Aid Study (NPSAS) and the Baccalaureate and Beyond (BB); (ii) the US Department of Education, and (iii) the Federal Reserve of New York, which runs the Consumer

Credit Panel (CCP) jointly with Equifax and collects information for over 40 million agents.

Table A2: Description of Main Business Variables

Variable	*Description*
Ownership share	Continuous measure for the share in firm's ownership by respondents.
Hours worked	Average number of hours per week devoted to the business.
Legal status	Categorical variable for the legal status of the firm. Categories are sole proprietorship, partnership, limited liability company or corporation.
Collateralized debt	Business finance collateralized by the owner using personal assets.
Employees	Number of employees working for the business of the respondent.
Gross sales	Gross sales receipt in the year before the time of the interview.
Profits	Total pre-tax net income in the year before the time of the interview.
Net worth	Value at which respondent could sell the business at the time of the interview. Should exclude business loans and include business assets (implements and materials too).
Business age	Survey year minus the year in which the business was started.
Business origin	Categorical variable for whether the business was "started", "bought", "inherited" or "joined" by the respondent.
Sector FE	It refers to the 1-digit industry code.

Table A3: Description of Student Loans Variables

Variable	*Description*
Number of loans	Total number of education loans. Possible range: 0 to 6. However, 99% of the sample considered has between 0 and 3 education loans.
Amount of loan	How much was borrowed, not counting the finance charges
Amount to be repaid	How much is still owed on the loan at the time of interview
Repayment rate	Amount to be repaid periodically until extinguishing the loan
Interest rate	Annual rate of interest charged on the loan
Year loan taken	Year respondent took out his/her loan
Year started repayment	Year respondent started making payments on his/her loan
On schedule	Categorical variable for whether the loan is being paid off ahead of schedule, behind schedule, or on schedule.
IBR	Whether the respondent is enrolled in a income based repayment plan

Table A5 reports figures related to the average and median amount of student debt, considering all borrowers. The average amount of student debt upon graduation and per graduate (as opposed to per borrower) was $18,650 in 2004 and $24,200 in 2011 according to NPSAS, similarly to estimates from SCF. In 1992, the average debt at graduation was 13,500 according to NPSAS, and 12,538 according to SCF. As of 2019, the total amount of student loans is reported to be worth 1.4 trillions of dollars in SCF, 1.6 trillions of dollars in NCES and 1.7 trillions of dollars in FRED.

One can also compare the SCF to other datasets under different dimensions, such as patterns in

Table A4: Student Loans in SCF and Other Sources: Part 1

	1989 – 1992		2007 – 2010	
	SCF	NPSAS	SCF	NPSAS (NCES)
% Borrowers in College Recipients	53%	55%	63%	68% (62%)
	2007 – 2010		2016 – 2019	
	SCF	Census	SCF	Census
% Borrowers in Civilian Population	11%	12%	14%	16%
% Borrowers in College Educated Households	30%	33%	36%	37%

Notes: When computing estimates in SCF, survey weights are used. Number of borrowers are from https://educationdata.org/student-loan-debt-statistics. Share of college educated households can be found at https://www.statista.com/statistics/184260/educational-attainment-in-the-us/. Civilian noninstitutional population is from https://www.bls.gov/emp/tables/civilian-noninstitutional-population.htm. Estimates from the National Postsecondary Aid Study (NPSAS) are from Hershbein and Hollenbeck (2015). Estimates from National Center Education Statistics (NCES) can be found at https://nces.ed.gov/programs/digest/d20/.

loan repayments and the distribution of loan balances. First, according to NCES, the cohort that entered student debt repayments in 2014 has shown a 12% default rate, compared to a 15% default rate computed using SCF and focusing on agents declaring that their student loan payments are "behind schedule". As reported by Brown et al. (2015) using CCP data on 40 millions individuals, 20% of borrowers still in repayment by 2004 were 90+ days late on their payments, against a 18% computed in SCF for the same year. Secondly, in 2014, the National Student Loan Data System (NSLDS) estimated that among all borrowers, 42% of them had balances in excess of $25K, 17% of them had more than $50K and 5% of them had more than $100K. Using SCF data, I can compute those shares to be 48%, 24% and 5% respectively in 2014. Going back in time instead, in 1992 only 8% and 2% of borrowers had more than $25K and $50K student debt balances respectively according to NSLDS, and such figures line up with those estimated in SCF (9% and 2%).

On the negative side, it has been argued that, in 2013, the SCF underestimated the share of debt held by the top quintile of the income distribution compared to what administrative data merged with sources from the US Department of Education seem to suggest (27% against 35%).[1]

A.2 Descriptive Statistics

In Figure A.1, I report the negative correlation between the average business ownership rate and the average student debt per person over time, considering loans with balances greater than 0 at the time of the interview for the sake of the computation. The graph controls for demographic characteristics such as gender, age, educational level, marital status, ethnicity and assets, and uses survey weights to ensure representativeness. Then, Figure A.2 breaks down the legal type of the businesses opened by college graduates with and without student loans. Possible categories are given by "sole-proprietorships", "partnerships", "corporations" (including C and S-corporations), and "limited liabilities companies". In the first two categories, the entrepreneurs have themselves unlimited liability for the business they run, either alone or with a partner. Both the second two

[1] See the full comparison here: https://www.brookings.edu/blog/up-front/2019/06/28/who-owes-the-most-student-debt/.

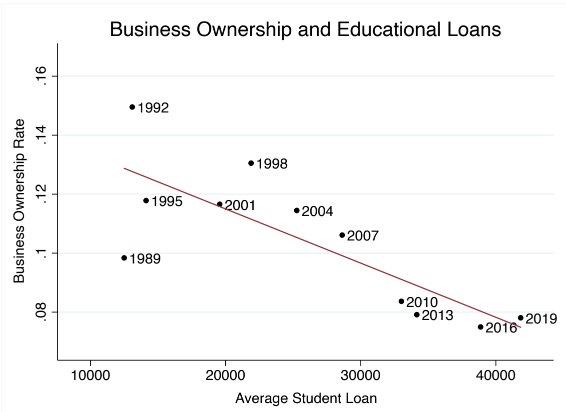
Table A5: Student Loans in SCF and Other Sources: Part 2

	2004			2010			2019		
	SCF	NPSAS	CCP	SCF	NPSAS	CCP	SCF	NCES	FRED
Avg. Amount	16,908	18,650	15,308	21,745	24,200	21,842	29,488	29,500	–
Median Amount	11,800	11,600	12,332	15,000	14,083	–	25,000	–	–

Notes: When computing estimates in SCF, survey weights are used and I winsorize data at the 99th percentile to exclude possible outliers and most likely misreported figures. Recent data on the total value of student debt is from the Federal Reserve Bank of St. Louis at `https://fred.stlouisfed.org/series/SLOAS`. Estimates from the National Postsecondary Aid Study (NPSAS) are from Hershbein and Hollenbeck (2015). For the *median amount* of 2010, I impute the value based on the growth rate of the median amount in NPSAS data over the 2004-2008 period. Estimates from National Center Education Statistics (NCES) can be found at `https://nces.ed.gov/programs/digest/d20/`.

categories provide limited liability protection, with the main difference being that a LLC is owned by one or more individuals, and a corporation is owned by its shareholders.

Figure A.1: Comparison over Time: 1989-2019



Moreover, in the right panel of Figure A.2, I report the average age (in years) of the firms started by entrepreneurs that have a college degree, distinguishing for whether they had to take student loans or not (I could alternatively focus as well on those still repaying their loans at the time of the survey interview). In the SCF, individuals can indicate whether the business they own and actively manage was either "bought", "started", "inherited" or "joined". In the right panel of Figure A.2, I consider entrepreneurs that started their own business and have the same educational attainment, and I find that owners who had to borrow for college run firms that are on average 5 years younger, suggesting a delay in the business funding year.

Finally, I analyse distributional properties of wages and profits for workers and entrepreneurs in the SCF data, pooling together all sample years and without conditioning on any control variable. Figure A.3 shows that, while the average and median values of wages and profits follow different patterns and growth trajectories over individuals' life-cycle, measures of relative volatilities stay virtually unchanged and stable, which justifies my modeling choices.

Figure A.2: Legal Status and Business Age

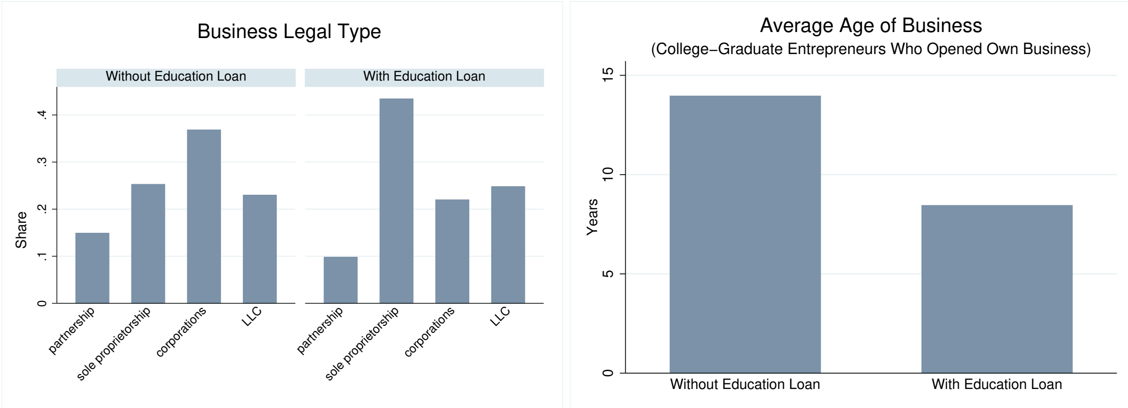
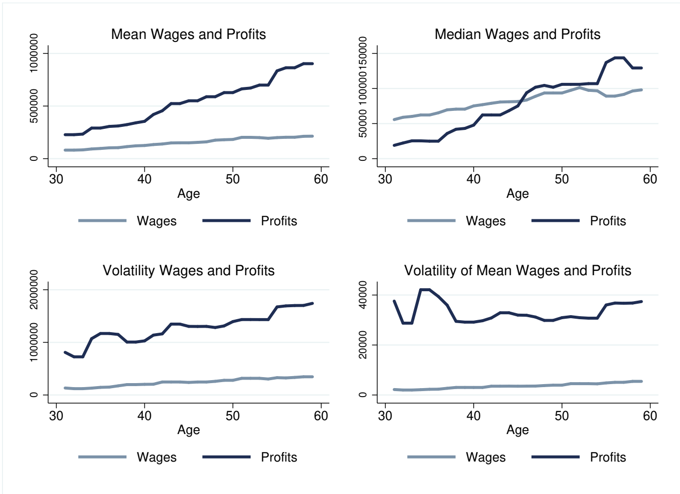


Figure A.3: Distributional Properties of Wages and Profits over the Life-Cycle



A.3 Additional Regression Results

To start, Table A6 reports that entrepreneurs with larger amounts of student debt (either considering the initial debt taken or the balance still to be repaid at the time of the interview) employ more personal collateral for their firms, comparing enterprises and owners of similar characteristics.

In Table A7, I provide a robustness check for the regressions estimated in Table 2, in which I control for respondents' net worth and isolate the correlation between student debt and business ownership beyond the impact that wealth and other outstanding loans can have on entrepreneurial entry. Note that, for this specific exercise, I use a different version of the SCF dataset, which contains summary variables for individuals' assets and liabilities. In particular, the website of the Federal Reserve provides access to an online tool called Summary Extract Public Data Files (SDA), from which I am able to construct a net worth variable defined as *Total Assets – Total Liabilities* and that excludes student loans. For instance, "total assets" merge together financial and non-financial wealth, including residences, vehicles, saving accounts, and any amount invested in mutual funds, stocks, pensions and bonds, among others. On the contrary, "total debt" includes principal residence debt (mortgages and HELOCs), lines of credit, debt for other residential property, credit card debt and installment loans, among others. Liabilities are to be considered as

Table A6: Collateral

	(1)	(2)	(3)	(4)
log(Original Student Debt Taken)	-0.0018	0.0168*		
	(0.0086)	(0.0088)		
Dummy(Have Loan)			0.1672*	
			(0.0873)	
log(Student Debt Still Owed)				0.0158*
				(0.0092)
Pre-College Controls	Y	Y	Y	Y
General Controls	N	Y	Y	Y
Firm Controls	N	Y	Y	Y
Survey Year FE	Y	Y	Y	Y
Industry FE	N	Y	Y	Y
Observations	40,085	39,401	39,401	39,401
R^2	0.0169	0.0846	0.0846	0.0846

Notes: Robust standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1. Survey weights are used. The dependent variable is a binary indicator = 1 if the individual is a business owner, and = 0 if not. Pre-College controls refer to agent's gender and ethnicity. General control variables are agents' education, age, marital and home-ownership status and personal wealth. Firm controls include profits, business size, legal type and individuals working hours. Robust to include also spousal income and the leverage ratio of the households instead of their asset positions.

outstanding loans, including the amount of student debt reported (up to six education loans).

The SDA dataset differs from the one used in my main analysis insofar as it does not contain all the variables from the SCF questionnaire. For this reason, here I define as a business owner any respondent that actively manages a business, and then estimate again Equation 1 with outstanding student debt and individuals' net worth as main regressors. I also control for both pre-determined variables (eg: gender and ethnicity) and contemporaneous ones (eg: age, educational attainment, number of kids, marital status). Finally, I include survey year FE and apply survey weights.

Table A7: Business Ownership (Summary Extract Public Data Files)

	(1)	(2)
log(Student Debt Still Owed)	-0.0034***	-0.0018***
	(0.0002)	(0.0002)
Pre-College Controls	Y	Y
General Controls	N	Y
Survey Year FE	N	Y
Observations	170,357	170,357
Pseudo-R^2	0.0218	0.0641

Notes: Estimates are average marginal effects. Robust standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1. Survey weights are used. The dependent variable is a binary indicator = 1 if the individual is a business owner. Pre-College controls refer to agent's gender and ethnicity. General controls are agents' net worth, education, age, marital status and number of kids. Robust to including individual and spousal income.

Back to the dataset used for the main analysis, Table A8 reports the same regression as in Columns (1)-(2) of Table 2 and Table 5 controlling for parental education, which is available only for the 2016 and 2019 surveys. In Table A9, I instead conduct a robustness check for the results in Table 2 without restricting the firm ownership share to be 100% in order for individuals to count as business owners. In the SCF sample of entrepreneurs, 74% of them hold the entire ownership of their business, while almost 25% of them have at least a 50% share of their business. The share of entrepreneurs owning less than 50% of their firm is hence smaller than 1%, and it is not likely to change the quality and extent of my results. Finally, Table A10 carries out again the analysis in Table 2 but focusing only on the largest education loan reported by survey respondents.

Table A8: Entrepreneurial Margins (Controlling for Parental Education)

	Ownership	Ownership	Loan Approval	Loan Approval
log(Original Student Debt Taken)	-0.0031***	-0.0013**	-0.0180***	-0.0138***
	(0.0004)	(0.0004)	(0.0051)	(0.0033)
Pre-College Controls	Y	Y	Y	Y
General Controls	N	Y	N	Y
Firm Controls	N	N	N	Y
Survey Year FE	N	Y	Y	Y
Observations	31,652	31,004	1,422	1,422
R^2	0.0475	0.0641	0.2164	0.6311

Notes: Estimates are average marginal effects. Robust standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1. Survey weights are used. In Columns (1)-(2), the dependent variable is a binary indicator = 1 if the individual is a business owner. In Columns (3)-(4), the dependent variable is a binary indicator = 1 if the business owner received a business loan over the 12 months previous to the survey interview. *Pre-College Controls* refer to agent's gender, ethnicity and parental education. *General Control* variables include agents' education level, age, marital status and home-ownership status, and income. Firm controls include size, business age, legal type and individuals working hours. Robust to including spousal income, the leverage or the assets of the households, and to using an income or wealth category by age and education instead of their personal income.

Table A9: Entrepreneurial Rates (No Ownership Share Restriction)

	(1)	(2)	(3)	(4)
log(Original Student Debt Taken)	-0.0028***	-0.0017***		
	(0.0002)	(0.0003)		
Dummy(Have Loan)			-0.0188***	
			(0.0024)	
log(Student Debt Still Owed)				-0.0017***
				(0.0003)
Pre-College Controls	Y	Y	Y	Y
General Controls	N	Y	Y	Y
Firm Controls	N	Y	Y	Y
Survey Year FE	Y	Y	Y	Y
Industry FE	N	Y	Y	Y
Observations	160,262	160,262	160,262	160,262
Pseudo-R^2	0.0383	0.0456	0.0457	0.0456

Notes: Estimates are average marginal effects. Robust standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1. Survey weights are used. The dependent variable is a binary indicator = 1 if the individual is a business owner. Pre-College controls refer to agent's gender and ethnicity (robust to including parental education, available only for the years 2016/2019). General control variables are agents' years, age, marital and home-ownership status and income. Robust to including spousal income, the leverage or the assets of the households, and to using an income or wealth category by age and education instead of their personal income.

To complement the analysis conducted in Table 4, Table A11 shows how outstanding student debt balances correlate with the likelihood of business ownership for individuals of different age categories. I control for both pre-determined variables (eg: gender and ethnicity) and contemporaneous ones in a sequential way (eg: income, educational attainment, marital and home-ownership status). Finally, I include survey year FE and apply survey weights. Note that the regression is estimated non-parametrically and shows that the negative correlation between the amount of student debt owed at the time of the survey and business ownership decreases as individuals age. This is in line with the economic intuition that the repayment of college borrowing should have a stronger

Table A10: Business Ownership, Largest Education Loan Only

	(1)	(2)	(3)	(4)
log(Original Student Debt Taken)	-0.0030***	-0.0019***		
	(0.0002)	(0.0002)		
Dummy(Have Loan)			-0.0166***	
			(0.0025)	
log(Student Debt Still Owed)				-0.0021***
				(0.0002)
Pre-College Controls	Y	Y	Y	Y
General Controls	N	Y	Y	Y
Survey Year FE	N	Y	Y	Y
Observations	170,302	170,302	170,302	170,302
Pseudo-R^2	0.0279	0.0554	0.0552	0.0594

Notes: Estimates are average marginal effects. Robust standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1. Survey weights are used. The dependent variable is a binary indicator = 1 if the individual is a business owner. Pre-College controls refer to agent's gender and ethnicity (robust to including parental education, only available in 2016/2019). General controls are agents' education, age, marital and home-ownership status and income. Robust to including spousal income, households' leverage or assets, to using either an income or wealth category by age and educational level instead of agents' personal income, and to considering owners with any given equity share.

impact on entrepreneurial margins at the beginning of individuals' working career.

Table A11: Business Ownership (Interaction Student Debt and Age)

	(1)	(2)
log(Student Debt Still Owed)	-0.0202***	-0.0303***
	(0.0039)	(0.0040)
log(Student Debt Still Owed) × 31-40yo	+0.0045	+0.0088**
	(0.0036)	(0.0035)
log(Student Debt Still Owed) × 41-50yo	+0.0078***	+0.0148***
	(0.0041)	(0.0039)
log(Student Debt Still Owed) × >50yo	+0.0180***	+0.0275***
	(0.0042)	(0.0040)
Pre-College Controls	Y	Y
General Controls	N	Y
Survey Year FE	N	Y
Observations	27,587	27,330
Pseudo-R^2	0.0290	0.0919

Notes: Estimates are average marginal effects. Robust standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1. Survey weights are used. The dependent variable is a binary indicator = 1 if the individual is a business owner. Pre-College controls refer to agent's gender and ethnicity. General controls are agents' income, education, marital and home-ownership status, and the (log) original amount of student loan individuals graduated with.

In Table A12, I report the estimates for the likelihood of applying for a firm loan, given a set of control variables and the presence and extent of student loans in the household's balance sheet. The probability of applying for business credit is estimated via the following probit regression:

$$Pr(Apply_{it} = 1) = F\left(\beta_0 + \beta_1 Student\ Loan_{it} + \delta'\mathbf{\Gamma}_i + \gamma'\mathbf{\Phi}_{it} + \alpha_t + \epsilon_{it}\right) \tag{10}$$

where the outcome variable $Apply$ is an indicator equal to 1 if the entrepreneur mentions to have applied for a business loan in the 12 months before the interview took place, and 0 otherwise. Controls and regressors are the same as for the specifications reported in Table 5. The initial amount of student loans taken for college education does not correlate with the probability of applying for business funding (see Columns (1)-(2)). A similar observation holds true when using as main regressor a dummy for whether the individual carries still student debt balances to repay at the time of the interview, as shown in Column (3). The total amount to be repaid is only mildly significant, but the size of the standard errors calls for caution in interpreting the result.

Table A12: Business Loan Applications

	(1)	(2)	(3)	(4)
log(Original Student Debt Taken)	-0.0006	0.0014		
	(0.0009)	(0.0009)		
Dummy(Have Loan)			0.0098	
			(0.0093)	
log(Student Debt Still Owed)				-0.0017*
				(0.0009)
Pre-College Controls	Y	Y	Y	Y
General Controls	N	Y	Y	Y
Firm Controls	N	Y	Y	Y
Survey Year FE	Y	Y	Y	Y
Industry FE	N	Y	Y	Y
Observations	20,017	19,693	19,693	19,693
Pseudo-R^2	0.0283	0.1155	0.1154	0.1156

Notes: Estimates are average marginal effects. Robust standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1. Survey weights are used. The dependent variable is a binary indicator = 1 if the individual is a business owner, and = 0 if not. Pre-College controls refer to agent's gender and ethnicity. General control variables are agents' education, age, marital and home-ownership status and personal wealth. Firm controls include profits, business size, legal type and individuals working hours. Robust to include also spousal income and the leverage ratio of the households instead of their asset positions.

In Table A13, I run alternative specifications for the regressions included in Table 6, where I have analysed the association between student loans and business outcomes such as size and gross sales. Differently from the specifications in the main text, here I use as main regressors either a dummy variable that signals the presence of pending student loans in the balance sheet of the households, or the actual amount still to be repaid as of the survey year t. In Table A14, I report instead the results from the regression in Equation 3, focusing on profits and business net worth. Then, in Table A15, I conduct robustness checks on these very same specifications using as main regressors either a dummy variable that signals the presence of pending student loans in the balance sheet of the households, or the actual amount still to be repaid as of the survey year t. All the results are consistent with the baseline regressions in the main text.

In Table A16, I run alternative specifications for the regressions on firms' profitability included in Table 7, using as main regressors either a dummy variable that signals the presence of pending student loans in the balance sheet of the households, or the actual amount still to be repaid as of the survey year t. The full set of controls is used. Results are consistent with the baseline specifications in the main text: entrepreneurs with student loans to repay tend to have between 6% and 12% higher profitability, depending on the specification. Furthermore, an increase of 1000$ in the amount of student debt still to be paid is associated with 4% to 9% higher business profitability.

In Table A17, I show that individuals' cognitive abilities are correlated with both higher amounts of grants and education loans. To this end, I use the US 1997 National Longitudinal Survey of Youth, which surveys and track a panel of households that were between 12 and 17 years old in

Table A13: Business Outcomes: Size and Gross Sales

	Employees	Employees	Sales	Sales
Dummy(Have Loan)	-18.5950***		-0.4475***	
	(1.7959)		(0.0474)	
log(Student Debt Still Owed)		-2.0644***		-0.0436***
		(0.1975)		(0.0051)
Pre-College Controls	Y	Y	Y	Y
General Controls	Y	Y	Y	Y
Firm Controls	Y	Y	Y	Y
Survey Year FE	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y
Observations	39,461	39,461	36,855	36,855
Pseudo-R^2	0.0339	0.0339	0.4059	0.4053

Notes: Robust standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1. Survey weights are used. The dependent variables are either the number of employees or $\log(Sales)$. Pre-College controls refer to agent's gender and ethnicity (robust to include parental education, only available in 2016/2019). General control variables are agents' education, age, marital and home-ownership status and personal wealth. Firm controls include business age, legal type and individuals working hours (and business size in Columns (3)-(4)). Robust to include also spousal income and the leverage ratio of the households instead of their asset positions.

Table A14: Business Outcomes: Profits and Net Worth

	Profits	Profits	Net Worth	Net Worth
log(Original Student Debt Taken)	-0.0376***	-0.0294***	-0.0660***	-0.0523***
	(0.0057)	(0.0052)	(0.0045)	(0.0036)
Pre-College Controls	Y	Y	Y	Y
General Controls	N	Y	Y	Y
Firm Controls	N	Y	Y	Y
Survey Year FE	Y	Y	Y	Y
Industry FE	N	Y	Y	Y
Observations	33,673	33,014	36,001	43,988
R^2	0.0658	0.3219	0.0787	0.3150

Notes: Robust standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1. Survey weights are used. The dependent variables are the $\log(Profits)$ and log(*Net Worth*) of businesses, as reported by entrepreneurs in the sample. Pre-College controls refer to agent's gender and ethnicity (robust to including parental education, only available in 2016/2019). General controls are agents' education, age, marital and home-ownership status and income. Firm controls include business size, age, legal type and individuals working hours. Robust to including spousal income, households' leverage or assets, to using either an income or wealth category by age and educational level instead of personal income, and to considering owners with any given equity share.

1997 and were followed since then. The survey means to be representative of the population, but I again make use of sample weights to further ensure representativeness. In terms of educational outcomes, the survey records the amount of grants and loans received by agents during college. Moreover, it reports the results to the Armed Services Vocational Aptitude Battery (CAT-ASVAB), which measures the respondents' skills in Arithmetic Reasoning, Electronics Information, Numerical Operations, Assembling Objects, General Science, Paragraph Comprehension, Auto Information, Mathematics Knowledge, Shop Information, Coding Speed, Mechanical Comprehension and Word Knowledge. The resulting estimates summarize the respondent's performance on each subtest on a scale that can be compared across respondents: a lower score indicates poorer performance, and a higher score indicates better performance This measure was included also in the previous 1979 National Longitudinal Survey of Youth and has been used by researches to proxy for households' underlying abilities (see for example Guvenen et al. (2020)). In the regressions that follow, I hence use the scores of respondents as a measure of cognitive abilities.

Table A15: Business Outcomes: Profits and Net Worth

	Profits	Profits	Net Worth	Net Worth
Dummy(Have Loan)	-0.3314***		-0.5395***	
	(0.0504)		(0.0356)	
log(Student Debt Still Owed)		-0.0306***		-0.0550***
		(0.0055)		(0.0038)
Pre-College Controls	Y	Y	Y	Y
General Controls	Y	Y	Y	Y
Firm Controls	Y	Y	Y	Y
Survey Year FE	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y
Observations	33,014	33,014	43,988	43,988
Pseudo-R^2	0.3224	0.3218	0.3157	0.3151

Notes: Robust standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1. Survey weights are used. The dependent variables are the log($Profits$) and log(*Net Worth*) of businesses, as reported by entrepreneurs in the sample. Pre-College controls refer to agent's gender and ethnicity (robust to including parental education, only available in 2016/2019). General controls are agents' education, age, marital and home-ownership status and income. Firm controls include business size, age, legal type and individuals working hours. Robust to including spousal income, households' leverage or assets, to using either an income or wealth category by age and educational level instead of personal income, and to considering owners with any given equity share.

Table A16: Business Outcomes: Profitability

	$\log\left(\frac{\text{Profits}}{Revenues}\right)$	$\log\left(\frac{\text{Profits}}{Revenues}\right)$	$\log\left(\frac{\text{Profits}}{CollDebt}\right)$	$\log\left(\frac{\text{Profits}}{CollDebt}\right)$
Dummy(Have Loan)	0.1227***		0.0579***	
	(0.0243)		(0.0172)	
log(Student Debt Still Owed)		0.0128***		0.0062***
		(0.0027)		(0.0017)
Pre-College Controls	Y	Y	Y	Y
General Controls	Y	Y	Y	Y
Firm Controls	Y	Y	Y	Y
Personal Wealth	Y	Y	Y	Y
Survey Year FE	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y
Observations	39,461	39,461	39,461	39,461
R^2	0.1415	0.1413	0.0575	0.0575

Notes: Estimates are average marginal effects. Robust standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1. Survey weights are used. Pre-College controls refer to agent's gender and ethnicity (robust to including parental education, only available in 2016/2019). General controls are agents' education, age, marital and home-ownership status and income. Firm controls include business age, legal type and individuals working hours. Robust to including spousal income, households' leverage or assets, to using either an income or wealth category by age and educational level instead of personal income, and to considering owners with any given equity share.

I control for college characteristics (eg: public vs private), and individuals' characteristics that were pre-determined to their college choices, such as their gender, ethnicity, parental education, family income and birthday year. Higher cognitive abilities correlate with higher amounts of grants, which are likely to capture students' access to merit-based aid, whereas they do not relate to the total amount of loans take out by respondents to finance college education. Moreover, I can check that higher cognitive skills do not predict a higher amount of grants compared to loans. This is consistent with the fact that grants for US universities typically cover a fifth of the total university tuition and are available only to individuals meeting specific background characteristics. Moreover, grants tend to be complemented by either borrowing or out-of-pocket contributions.

Using again the panel of respondents from the NLSY97, Table A18 shows that student debt is

Table A17: Educational Outcomes in NLSY97

	Difference Grants vs Loans	Total Loans	Total Grants
Cognitive Skills	0.0025	0.0007	0.0031**
	(0.0019)	(0.0014)	(0.0013)
Controls	Y	Y	Y
Observations	4,107	4,873	5,765
R^2	0.1005	0.0776	0.1317

Notes: Robust standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1. Survey weights are used. Controls include agent's gender and ethnicity, parental education, age, college type, college tuition, full-time vs part-time college attendance, and family income. Robust to the inclusion of *Cognitive Skills* as the only main regressor.

negatively associated with the likelihood of owning a firm even after controlling for individuals' cognitive skills. This strengthens the idea that the negative correlation between student debt and entrepreneurial outcomes found in the SCF is not driven by a group of particularly low-skilled households who happen to have taken out large amounts of education loans. In particular, I run the following set of probit regressions:

$$Pr(BusOwn_{it} = 1) = F\left(\beta_0 + \beta_1 Student\ Loan_i * Cognitive\ Skills_i + \delta' \mathbf{\Gamma}_i + \gamma' \mathbf{\Phi}_{it} + \alpha_t + \epsilon_{it}\right) \quad (11)$$

Table A18: Business Outcomes in NSLY97

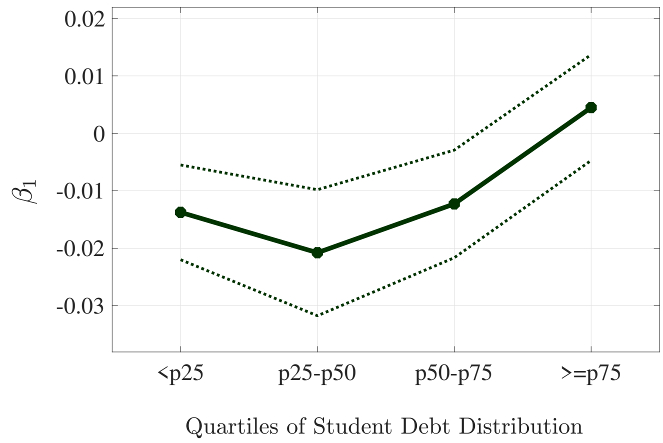
	Ownership	Ownership	Ownership	Ownership
Dummy(Have Loan)	-0.0231***	-0.0279***		
	(0.0053)	(0.0072)		
Amount Taken			-0.0189***	-0.0189
			(0.0061)	(0.0122)
Pre-College Controls	Y	Y	Y	Y
General Controls	N	Y	Y	Y
Year FE	Y	Y	Y	Y
Observations	48,345	28,688	8,354	8,354
R^2	0.0225	0.0242	0.0411	0.0411

Notes: Estimates are average marginal effects. Robust standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1. Survey weights are used. Pre-College controls refer to agent's gender and ethnicity, parental education and income and birthday year. General control variables are agents' marital status, region of residency and assets. Column (4) clusters standard errors at the individual level and has a *p-value*=0.12.

where Y_{it} is a dummy signaling whether the respondent is an active business owner or not. I include self-employed individuals as I cannot define firm owners in the exact same way I did for SCF, namely focusing on ownership shares and presence of salaried workers. I include both controls that were pre-determined to the choice of education, as in Table A17, and contemporaneous control variables such as their region, marital status and wealth. Results are shown in Table A18 for the main regressors of interest, which are (i) an indicator for whether the household took out student debt, and (ii) the original amount of education loans contracted.

Finally, Figure A.4 shows the different elasticities of business ownership to outstanding student loans by quartiles in the distribution of total student debt taken out for college purposes (in dashed are the 95% confidence intervals). Regressions include controls as in Column (4) of Table 2 and survey weights, but condition on individuals that contracted education loans, as opposed to use the entire sample of SCF respondents. The purpose is to illustrate heterogeneity in the negative association between college debt and the extensive margin of entrepreneurship, and to show that

Figure A.4: Elasticity of Business Ownership to Student Loans



results are not driven by individuals in the top percentiles of the student debt distribution.

B Bankruptcy Reform

Instead of an RDD, to quantify the impact of the 1998 bankruptcy reform on the extensive margin of entrepreneurship I can estimate a diff-in-diff probit regression of the following form:

$$Pr(BusOwner_{it} = 1) = F\left(\beta_0 + \beta_1 Post_{it} + \beta_2 Reform_{it} + \beta_3 Post_{it} \times Treated_{it} + \gamma' \mathbf{\Phi}_{it} + \alpha_t + \epsilon_{it}\right)$$

where $BusOwner$ is a binary variable equal to 1 if individuals are entrepreneurs at the time of the survey, and to 0 if they are not. The regressor $Post_{it}$ captures the difference in business ownership rates before and after the 7th year of repayment, while $Treated_{it}$ is an indicator equal to 1 if individuals fall in the treated group and 0 if they belong to the control group. I consider three cases: in the first regression, the treatment group includes agents that started repaying their debt between 1992 and 1997, and the control group includes those that started repaying in or before 1991. In the second case, the treatment group is composed of individuals that started their repayment in or before 1991 but had still not finished repaying their education loans, while the control group contains households that had finished their repayment period by the time the reform stroke. Finally, a third set of regressions compares individuals who started repaying their loans between 1992 and 1997 to a control group composed of those who started repaying after the 1998 reform took place.

In all three cases, the coefficient of interest is β_3, which captures the differential likelihood of transitioning into entrepreneurship for individuals that were subject to the reform and after their 7th year in student debt repayment. I then include a set of controls $\mathbf{\Phi}$, which capture factors pre-determined to the choice of taking on student loans and also include variables recorded at the time of the survey that were not pre-determined at the time in which the individuals made their student loans choices, such as their age, educational level, marital and home-ownership status, and personal wealth. All regressions include survey year fixed effects (α_t) and use survey weights.

Columns (1) to (2) in Table B19 report the results of the first set of regressions, comparing individuals who started repaying their student loans before or after 1991. The inclusion of controls that are not pre-determined by the time the loan was taken does not alter the estimates: households who did not reach the 7th year into repayment by 1998 are 13% less likely to have become entrepreneurs.

Table B19: Business Ownership

	(1)	(2)	(3)	(4)
Post×Reform	-0.1190***	-0.1348***	-0.2363***	0.0118
	(0.0269)	(0.0299)	(0.0492)	(0.0186)
Pre-College Controls	Y	Y	Y	Y
General Controls	N	Y	Y	Y
Personal Wealth	N	Y	Y	Y
Survey Year FE	Y	Y	Y	Y
Observations	4,398	4,398	3,421	17,756
Pseudo-R^2	0.0390	0.0644	0.0772	0.0213

Notes: Estimates are average marginal effects. Robust standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1. Survey weights are used. The dependent variable is a binary indicator = 1 if the individual is a business owner, and = 0 if not. Pre-College controls refer to agent's gender and ethnicity. General control variables are agents' education, age, marital and home-ownership status. Robust to include also spousal income and the leverage ratio of the households instead of their asset positions.

Table B20: Business Performance: $\frac{\text{Profits}}{\text{Revenues}}$

	(1)	(2)	(3)
Post×Reform	0.0603**	0.16138***	-0.0065
	(0.0311)	(0.0485)	(0.0238)
Pre-College Controls	Y	Y	Y
General Controls	Y	Y	Y
Personal Wealth	Y	Y	Y
Survey Year FE	Y	Y	Y
Observations	4,398	3,421	17,756
R^2	0.0783	0.0832	0.0884

Notes: Estimates are average marginal effects. Robust standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1. Survey weights are used. The dependent variable is a binary indicator = 1 if the individual is a business owner, and = 0 if not. Pre-College controls refer to agent's gender and ethnicity. General control variables are agents' education, age, marital and home-ownership status. Robust to include also spousal income and the leverage ratio of the households instead of their asset positions.

Moreover, Column (3) focuses on agents that started repaying before 1991 but compares those who had and had not finishing paying their loans by 1998. Interestingly, the sign and magnitude of the estimated coefficient illustrates that those who were on time to declare bankruptcy before the reform took place, but lost such opportunity, are less likely to become entrepreneurs compared to those who were completely done paying by 1998. Since the regressions control for survey year fixed effects, the results are unlikely to be due to a declining time trend in business entry only. This is further confirmed by the estimate in Column (4), which shows that college graduates who started repaying between 1992 and 1997 are not less likely to become entrepreneurs compared to the new cohorts who started repaying after the reform took place.

Furthermore, as reported in Table B20, being subject to the reform and hence not being able to discharge student loans in bankruptcy has a positive effect on the profit margin of treated entrepreneurs, consistent with a phenomenon of stricter selection into the entrepreneurial pool. Once again, the effect primarily regards individuals who started repaying after 1991 but before the reform took place, and agents who started repaying before 1991 but did not finish repaying their loans by 1998. The results therefore suggest an effect of the 1998 reform to student loans bankruptcy availability also on the *intensive* margin of entrepreneurship for treated cohorts.

In Table B21, RDD regressions show that being past the 7th year of education loan repayment

Table B21: Outstanding Student Debt and Bankruptcy

	(1)	(2)
Past 7th Repayment Year by 1998	0.0997**	0.0788*
	(0.0449)	(0.0446)
Pre-College Controls	Y	Y
General Controls	N	Y
Survey Year FE	Y	Y
Observations	2,167	2,142
R^2	0.5374	0.5533
F-Statistic	477.70	266.10

Notes: Estimates are average marginal effects. Robust standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1. Survey weights are used. The dependent variable is the outstanding student debt balances reported by individuals interviewed up to 10 years after the 1998 reform that were within 10 years into their loans repayment by 1998. Pre-College controls refer to agent's gender, cohort year, and ethnicity. General control variables are agents' education, loan size, marital and home-ownership status, income and age. Robust to include also spousal income and the leverage ratio of the households instead of their asset positions.

correlates with lower outstanding student debt balances for cohorts that had the bankruptcy option available. I both control for factors that pre-determined to the choice of college and student debt as well as a battery of subsequent controls that can be contemporaneous to the choice of becoming entrepreneurs. To avoid the confounding effect of few outliers that repaid their loans for more than 2 decades, I consider respondents that started repaying before 1998 are were surveyed at most 10 years after the 1998 bankruptcy reform was passed (i.e. up to the 2009 survey). In particular, after controlling for age effects (including interactions as well), being past the 7th year of repayment is shown to be associated with a lower amount of outstanding student debt balances.

In Table B22, I document that being past the 7th year of educational loan repayment correlates with the likelihood of transitioning into entrepreneurship only for cohorts that had the bankruptcy option available. I both control for factors that pre-determined to the choice of college and student debt as well as a battery of subsequent controls that can be contemporaneous to the choice of becoming entrepreneurs. In particular, after controlling for age effects, being past the 7th year of repayment for recent cohorts does not matter anymore, but used to matter for cohorts that had the possibility to declare bankruptcy on their student debt after 7 years into full repayment.

Table B22: Business Ownership

	(1) After 1991	(2) Before 1991	(3) After 1991	(4) Before 1991
Past 7th Year	-0.0269	0.5745***	-0.0286	0.5773***
	(0.0247)	(0.0971)	(0.0251)	(0.1121)
Pre-College Controls	Y	Y	Y	Y
General Controls	N	N	Y	Y
Personal Wealth	N	N	Y	Y
Survey Year FE	N	N	Y	Y
Observations	17,751	1,768	17,751	1,768
Pseudo-R^2	0.0141	0.0569	0.0232	0.0973

Notes: Estimates are average marginal effects. Robust standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1. Survey weights are used. The dependent variable is a binary indicator = 1 if the individual is a business owner, and = 0 if not. Pre-College controls refer to agent's gender, cohort year, and ethnicity. General control variables are agents' education, loan size, marital and home-ownership status. Robust to include also spousal income and the leverage ratio of the households instead of their asset positions.

In Table B23, I run a similar set of regressions as in Table B22 to show that, before the 1998 reform took place, being past the 9th or 10th (or higher) repayment year cutoffs had no relationship with the likelihood of transitioning into entrepreneurship. The relevant repayment year cutoff was the 7th or the 8th one, suggesting that probably most of bankruptcy discharges were happening as soon as agents were past the 7th year into repayment had had legal access to the bankruptcy option. Moreover, I also check that these cutoffs are no longer significantly associated with the likelihood of transitioning into entrepreneurship for cohorts that started repaying their loans after 1991 and hence did not have any bankruptcy regime available (all results available upon request).

Table B23: Business Ownership

	(1) Before 1991	(2) Before 1991	(3) Before 1991
Past 8th Year	0.3375**		
	(0.1233)		
Past 9th Year		-0.0794	
		(0.1299)	
Past 10th Year			-0.1580
			(0.1194)
Pre-College Controls	Y	Y	Y
General Controls	Y	Y	Y
Personal Wealth	Y	Y	Y
Survey Year FE	Y	Y	Y
Observations	1,768	1,768	1,768
Pseudo-R^2	0.0818	0.0741	0.0956

Notes: Estimates are average marginal effects. Robust standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1. Survey weights are used. The dependent variable is a binary indicator = 1 if the individual is a business owner, and = 0 if not. Pre-College controls refer to agent's gender, cohort year, and ethnicity. General control variables are agents' education, loan size, marital and home-ownership status. Robust to include also spousal income and the leverage ratio of the households instead of their asset positions.

Figure B.1 shows the discontinuity in the likelihood of becoming an entrepreneur by repayment year, considering individuals that were repaying their student loans around the time of the 1998 bankruptcy reform. The underlying regression is estimated using the - rdplot - package from Calonico et al. (2015), using a polynomial fit of order 1, survey weights, no covariates and the default triangular kernel function to smooth observations.

Figure B.1: RDD Estimates

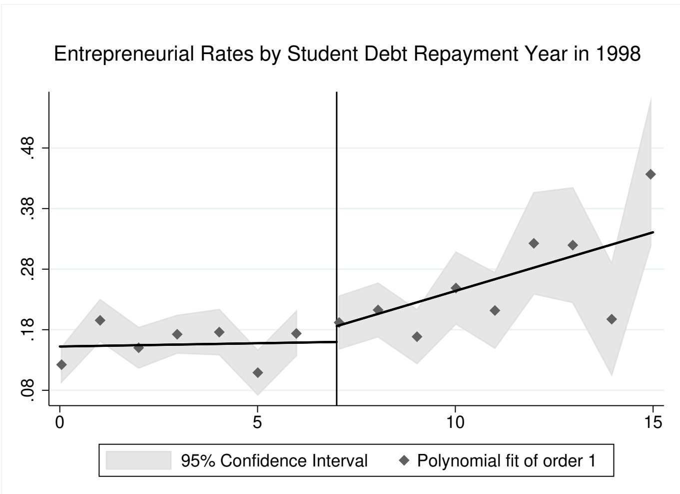


Table B24: Covariates and Treatment Effect

	Gender	Ethnicity	Marital Status	Assets	Amt Ed.Loan	Age
Subject to Reform	-0.0206	0.0486	0.0239	0.3265	-0.0811	0.3313
	(0.0175)	(0.0313)	(0.0221)	(0.3283)	(0.0924)	(0.7850)

Notes: Robust standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1. Survey weights are used.

Furthermore, Table B24 checks that the main covariates included in Table 13 and Table 14 do not correlate with being in the treated or control group. To do that, I perform simple OLS regressions to assess the correlation between the covariates and the indicator function for whether households are in the treated group, considering individuals that were in a window of 3 years from the 7th repayment year cutoff. Note that Columns (3)-(6) also include the running variable as in Equation 9 to control for time (or cohort) effects that could otherwise confound the estimates.[2]

Table B25: Business Ownership

	(1) Non-Affected 2-Y Bandwidth	(2) Non-Affected 4-Y Bandwidth	(3) Affected 2-Y Bandwidth	(4) Affected 4-Y Bandwidth
Subject to Reform	0.0113	0.0111	0.0286	0.0075
	(0.0471)	(0.0431)	(0.0280)	(0.0261)
Pre-College Controls	Y	Y	Y	Y
General Controls	Y	Y	Y	Y
Personal Wealth	Y	Y	Y	Y
Survey Year FE	Y	Y	Y	Y
Observations	1,310	1,531	2,133	2,660
Pseudo-R^2	0.0918	0.0755	0.0680	0.0538

Notes: Estimates are average marginal effects. Robust standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1. Survey weights are used. The dependent variable is a binary indicator = 1 if the individual is a business owner, and = 0 if not. Pre-College controls refer to agent's gender, cohort year, and ethnicity. General control variables are agents' education, loan size, marital and home-ownership status. Robust to include also spousal income and the leverage ratio of the households instead of their asset positions.

Finally, Table B25 conducts placebo tests to assess the validity of the RDD regressions in Table 13. In Columns (1) and (2) I include individuals that were theoretically past the 7th year repayment cutoff. My running variable counts the distance (in years) from a fictitious 9th repayment year cutoff and hence compare cohorts that, for example, started repaying their loans between 1988 and 1991 and cohorts that started repaying between 1984 and 1987. In Columns (3) and (4) I include individuals that were theoretically all below the 7th year repayment cutoff and hence are all treated by the reform in 1998. My running variable counts the distance from a fictitious 4th repayment year cutoff and hence compare cohorts that, for example, started repay their loans between 1992 and 1994 and cohorts that started repaying between 1995 and 1997. I consider 2 years and 4 years bandwidths to show that the results are not driven by the choice of the window around the cutoff year of interest. I show regression outcomes for the full set of control variables, but results are robust to the inclusion of pre-determined controls only (available upon request).

[2]Within a window around the cutoff point, individuals belong to different cohorts. What this means is that, for example, they would be likely to have accumulated more or less assets, or to be a couple of years younger or older.

C Model Specifications

C.1 Time Transformation: the Education Stage

In Section 3, I have developed a life-cycle model of education and occupational choices, where the reference period for calibration purposes is a year. Since my framework has abstracted from other dynamic choices during the education stage of individuals' life after they decide to enroll in university (eg: college dropout), and there is full commitment to graduate, I have assumed that agents spend only 1 period in college to save on notation and simplify the analysis. Yet, I have to ensure the time-consistency across all three stages in life (i.e: education, working life and retirement) for the correct quantification and calibration of the model. In particular, since in real-life agents spend 4 years in college, the value function for young adults that decide to enroll in university in their first year of their life $t = 1$ (or T_{edu}) is given by:

$$V^c(a_t, z_t, age_t) = \max_{a_{t+4}, d_{edu,t}, c_t} \left\{ \sum_{t=1}^{4} \beta^{t-1} u(c_t) + (\beta \theta_{age_t})^4 \int W^c(a_{t+4}, z_{t+4}, d_{t+4}, age_{t+4}) d\Xi(z_{t+4}|z_t) \right\}$$
$$\text{s.t.}: 4 * c_t + a_{t+4} = (1 + r_t)^4 a_t + 4 * (d_{edu,t} - \chi_t)$$
$$\text{and}: a_{t+4} \geq 0, \quad c_t \geq 0, \quad 0 \leq d_{edu,t} \leq \underline{d}$$

where I assume that individuals maintain the same profile of consumption across the 4 years spent in university, they pay the same yearly tuition – net of grants, as explained in previous sections – and choose the same yearly amount of college loans to pay for it. Their assets a_t are capitalized for 4 periods, and their future value W^c is discounted at the rate β^4. In a similar spirit, the value function for young adults that do not enroll in college in their first year of their life $t = 1$ (or T_{edu}), enter directly the labor markets and choose between being workers or entrepreneurs is given by:

$$V^{nc}(a_t, z_t, age_t) = \max_{a_{t+4}, c_t} \left\{ \sum_{t=1}^{4} \beta^{t-1} u(c_t) + (\beta \theta_{age_t})^4 \int W^{nc}(a_{t+4}, z_{t+4}, age_{t+4}) d\Xi(z_{t+4}|z_t) \right\}$$
$$\text{s.t.}: c_t + a_{t+4} = (1 + r_t)^4 a_t + (1 - \tau) \max\{\pi(a_t, z_t, age_t; r_t, w_t); \tilde{\omega}_{age_t}\}$$
$$\text{and}: a_{t+4} \geq 0, \quad c_t \geq 0, \quad k_t \leq \lambda a_t$$

C.2 Introducing a Corporate Sector

In an alternative version of the model, I include an unconstrained sector that contributes to total production in equilibrium. I do this to check that my results are not driven by the fact that my baseline economy has only one productive sector, in which entrepreneurs are constrained and in which outstanding student loans reduce the collateral that can be pledged to rent capital. Note that, in models à la Cagetti and De Nardi (2006) that include both entrepreneurial and non-entrepreneurial firms, it is often assumed that entrepreneurs produce using only capital, so that the size of the non-entrepreneurial sector is pinned down by the measure of workers in equilibrium (i.e: the share of the population who is not entrepreneurs). To remain close to the assumptions of the framework laid down in Section 3, I augment my economy with a corporate sector where unconstrained firms have all the same productivity and produce using capital and labor. Since corporate firms rent capital and labor as well, to obtain a well-defined measure of the unconstrained sector I have to assume that corporate firms operate according to a decreasing returns to scale technology with span of control parameter ν. Their production function is given by:

$$f(k,l) = A(k^{\alpha}l^{1-\alpha})^{1-\nu}, \quad \text{with} \qquad 0 < 1-\nu < 1$$

Table C26: Alternative Calibration w/ a Corporate Sector

Fitted	Value	Description	Moment	Model	Data
β	0.99	Discount factor	Interest rate	0.05	0.04
χ	1.00	College tuition	Educational rate	0.31	0.35
σ_a	3.50	Dispersion initial wealth	Top10 wealth share	0.69	0.70
ρ_{az}	0.25	Correlation initial (a,z)	Inter-generational earnings	0.30	0.28
ν	0.78	Entrepreneurs span of control	Top10 income share	0.43	0.45
A	1.475	Corporate productivity	Share corporate employment	0.30	0.30
σ_ϵ	0.305	St deviation prod shocks	Top25 employment share	0.57	0.65
ρ_z	0.92	Persistence entrep prod	Serial correlation revenues	0.84	0.80
λ	1.65	Financial constraint 1	Avg. corporate debt/GDP	0.27	0.35
η	0.125	Financial constraint 2	Δ Entr rates w/ – w/o Sloans	5pp	5pp
ζ_1^c	0.0573	Trend income growth (college)	Income growth year 0 - 30	0.84	0.86
ζ_2^c	0.0012	Curv. income growth (college)	Income growth year 30 - 40	0.07	0.05
ζ_1^{nc}	0.0310	Trend income growth (no coll)	Income growth year 0 - 30	0.48	0.48
ζ_2^{nc}	0.0004	Curv. income growth (no coll)	Income growth year 30 - 40	0.08	0.10

In each period t, corporate firms rent capital and hire labor at the equilibrium input prices $r_t + \delta$ and w_t, always determined in GE. Their profits are then distributed lump-sum to all households in the economy. In essence, corporate firms will differ from entrepreneurial businesses in two dimensions. First, their productivity A will be allowed to differ from the one of the entrepreneurial sector to reflect size differences across entrepreneurial businesses and corporations. Second, corporate firms will not face a borrowing limit when renting capital using financial markets. Thus, I modify my calibration strategy to be so that the value assigned to A imply that the share of employment of the corporate sector is 30%, as estimated for the US based on Compustat firms (see Davis et al. (2006)). Results from the estimation are presented in Table C26.

There are three main differences in the calibrated values of this extended model version with respect to the baseline case. The first one, is a 20% decrease in the parameter η that governs the student debt-related entrepreneurial borrowing constraint. Since the presence of another productive sector increases the demand for capital and labor, the GE wage and interest rate increase, further discouraging indebted college graduates from entering entrepreneurship. Secondly, the discount factor β has to rise to compensate for the upwards pressure on the equilibrium interest rate caused by the increase in the demand for capital in the economy. Moreover, a higher GE wage induces young adults at the margin to select out of college, which implies a slightly lower calibrated value for the college tuition κ in order to match the average college attainment rate in the US over the last decade (parallel to that, the value for the parameter ϕ_1 governing the extent of need-based grants increases by more than 20% to match the average share of tuition covered by means-tested scholarship). Finally, the fit of untargeted moments is close to the one of the baseline model: yet, since the presence of another productive sector increases the demand for capital and labor, the GE wage and interest rate increases, which lowers entrepreneurship by roughly 1 p.p. with respect to the baseline economy for both college and non-college graduates, with and without loans. Their respective share within the entrepreneurial sample stays instead relatively the same.

C.3 Introducing Student Debt Forbearance

As of today, roughly 20% of outstanding education loans are reported to be in deferment or forbearance, two available options for borrowers who are (currently) unable to pay back their debt, but intend to in the future.[3] The main differences between these two options for pausing student debt payments regard the average length of the programs, their qualifying requirements and the accrue of interest rates. In particular, deferment typically can last three years, while forbearance is granted for maximum 12 months for at most 3 times. To qualify for deferment, agents have to prove they are enrolled in school at least half time, or they are facing financial hardships, such as being unemployed or undergoing medical treatment for example. Instead, a specific qualifying event is usually not necessary to file for forbearance. Finally, under deferment, interest does not accrue on subsidized federal student loans and Perkins loans, while interest accrues on all types of loans under forbearance. Since 90% of borrowers in deferment are those who enroll in post-graduate schools, not included in my model, I keep my focus on student debt forbearance, especially given the relevance that forbearance has played in the recent pandemic years.[4]

Table C27: Alternative Calibration w/ a Student Debt Forbearance

Fitted	Value	Description	Moment	Model	Data
β	0.98	Discount factor	Interest rate	0.04	0.04
χ	1.25	College tuition	Educational rate	0.37	0.35
σ_a	3.50	Dispersion initial wealth	Top10 wealth share	0.69	0.70
ρ_{az}	0.25	Correlation initial (a, z)	Inter-generational earnings	0.31	0.28
ν	0.79	Entrepreneurs span of control	Top10 income share	0.45	0.45
μ	0.875	Prob delaying payments	Share Sdebt in forbearance	0.20	0.20
σ_ϵ	0.305	St deviation prod shocks	Top25 employment share	0.63	0.65
ρ_z	0.92	Persistence entrep prod	Serial correlation revenues	0.84	0.80
λ	1.65	Financial constraint 1	Avg. corporate debt/GDP	0.30	0.35
η	0.11	Financial constraint 2	Δ Entr rates w/ – w/o Sloans	5pp	5pp
ζ_1^c	0.0573	Trend income growth (college)	Income growth year 0 - 30	0.84	0.86
ζ_2^c	0.0012	Curv. income growth (college)	Income growth year 30 - 40	0.07	0.05
ζ_1^{nc}	0.031	Trend income growth (no coll)	Income growth year 0 - 30	0.48	0.48
ζ_2^{nc}	0.0004	Curv. income growth (no coll)	Income growth year 30 - 40	0.08	0.10

I introduce forbearance in the baseline model in a stylized way, by assuming that college borrowers are subject to a "forbearance" shock with probability μ, which allows them to stop repayments

[3]See `https://www.experian.com/blogs/ask-experian/research/student-loan-debt-and-repayment/`.

[4]For more information see `https://educationdata.org/deferment-vs-forbearance-student-loan`.

for a year. Although the decision to apply for forbearance is surely an endogenous choice, such simplifying assumption keeps the model tractable and allows me to qualitatively study the implications of student debt forbearance for my framework. Since, according to the current US legislation, student debt forbearance doe not impact individuals' credit scores, I assume that the only cost of forbearance is the interest rate accrued during the pause from repayment, which is capitalized on top of individuals' outstanding balances. Note that, since payments are suspended and interest is capitalized during the period spent in forbearance, the subsequent amount individuals have to pay by the end of their repayment term increases after (any) episode of forbearance. Finally, I calibrate μ such that the average time individuals spend in forbearance is 1.75 years.[5] Results from the calibration exercise for the internally fitted parameters are reported in Table C27.

Externally fixed parameters are by default kept at their baseline values. The main difference in the quantification of this model extension is that the value of η – the student debt-related entrepreneurial borrowing constraint – inferred through the calibration procedure is roughly 25% lower than in the baseline economy. This is consistent with the fact that, under forbearance, agents can pause their yearly payments and hence do not have to serve on their student debt obligations, which may increase the amount of capital they are able to rent as entrepreneurs through collateral pledgeability.[6] No other cost or credit reduction is in place after forbearance, and the higher repayment amounts individuals have to disburse after an episode of forbearance on average hit them later on in their life-cycle, when they have already accumulated savings as a buffer. As a final remark, note that the estimation of the model predicts a slightly higher value for the span of control parameter: this may indicative of the fact that, as forbearance allows entrepreneurs to rent higher capital, this counterfactual economy may in principle register an increase in entrepreneurship, which hence is partially counterbalanced by a decrease in the extent of entrepreneurial profits.

[5]See the report at `https://www.gao.gov/products/gao-18-163`.

[6]Interestingly, while almost all untargeted moments stay roughly the same in this model extension, the share of student borrowers decreases by 15 p.p. with respect to the baseline model (the education rate is instead targeted). This suggests that, the higher expected cost of taking out loans to finance college, given by the probability of pausing repayments and accumulate higher interest rates, discourage some borrowers at the margin from getting student debt.

Part II

Firms Dynamics in a New Keynesian Model

Andrea Chiavari*, Marta Morazzoni† and Danila Smirnov‡

1 Introduction

Far for being a closed topic of investigation, the discussion around the cyclicality of the aggregate markup and its response to monetary policy shocks still fosters a significant volume of macroeconomic research. Parallel to that, recent contributions have brought attention on companies' heterogeneous market power, as the availability of firm-level datasets has made it easier to estimate markups from balance sheet data. However, the empirical evidence of the heterogeneity in the behavior of firm-level markups after interest rate movements is scarce, and any related quantitative analysis has not yet been provided. This project is a first step towards filling this gap. In particular, we document crucial differences in the response of markups to monetary policy shocks by firm age, and assess their macroeconomic relevance into a novel New Keynesian framework enriched with firms' heterogeneity, demand accumulation and endogenous markups that evolve along firms' life-cycle.

We begin by estimating the behavior of markups at the company level conditional on interest rate movements, and document a significant degree of heterogeneity across old and young firms. Combining together quarterly data from Compustat with two different and exogenously identified series of monetary policy (hereafter MP) shocks for the US, we employ state-of-the-art local projection techniques to establish that the markups of firms above the median age respond more countercyclically to negative MP shocks, while for young firms the response is either mildly procyclical or insignificant. Controlling for commonly-used measures of aggregate economic activity and horse-racing our regression specifications with other firm-level characteristics, we are able to confirm that corporate age in particular influences the differential trajectory of markups upon a negative change in the interest rate. Moreover, we provide evidence that this result could indeed be related to a latent process of demand (or customer base) accumulation, for which dominant firms that are more established in their markets may have to change by less their prices in response to MP shocks, thereby leading to the stronger countercyclical response in old firms' markups documented in the data.

Next, we embed our findings into a New Keynesian (NK) framework that we enrich with firm heterogeneity, demand accumulation and endogenous markups. The model features imperfect competition among heterogeneous intermediate firms that produce using labor and choose prices to maximize profits subject to price adjustment costs à la Rotemberg. New firms can entry every

*Department of Economics, University of Oxford. Email: andrea.chiavari@economics.ox.ac.uk

†Department of Economics and Business, Universitat Pompeu Fabra. Email: marta.morazzoni@upf.edu. We are grateful to our advisors Isaac Baley, Davide Debortoli and Edouard Schaal for their invaluable support. We particularly thank Andrea Caggese, Jan Eeckhout, Andrea Fabiani, Jordi Galí, Basile Grassi, Priit Jeenas, Matthias Meier, Isabelle Mejéan, as well as all the participants at the CREI Macro Lunch, the 2022 "Conference on Individual Risk and the Macroeconomy" at Sciences-Po , and the 2022 Barcelona Summer Forum. All errors are our own.

‡Department of Economics and Business, Universitat Pompeu Fabra. Email: danila.smirnov@upf.edu

period, while the exit of incumbent firms is exogenous. Our framework presents therefore two main characteristics: on the one hand, intermediate firms face a process of demand accumulation, characterized by some persistence and idiosyncratic shocks, along with a long-run mean that allows for the demand faced by companies to increase along their life-cycle. On the other hand, we assume that the final good producer combines together the intermediate inputs by means of a Kimball aggregator. As in Klenow and Willis (2016) for example, this specific choice introduces in a tractable way endogenous markups in the model, as the elasticity of substitution across intermediate goods become decreasing in their relative quantity. Dominant firms will face lower elasticities of demand and, since demand is accumulated with age, older businesses will hence be able to charge higher markups.

The model is then calibrated on the US economy, following standard strategies in the literature and making use of the richness of Compustat data. In particular, the validation analysis shows that our quantitative framework is able to replicate several untargeted features of the data, such as the increasing profile of markups along the life-cycle of the firms, the fat right tail in the distribution of markups, and the growth rates of sales and employment. Moreover, the model can get realistic steady state distributions of businesses and employment shares by firm age, and replicate the elasticity of wages to firms' sales shares that we estimate in the data. This latter moment is tightly linked to the fact that dominant (and hence old) companies can increase their profits by cutting quantities and raising prices, thereby suppressing labor demand and hence wages in equilibrium.

Importantly, our NK framework enriched with firm heterogeneity, demand accumulation and endogenous markups can deliver the differential response of markups by firm age that we document in Compustat. As previously mentioned, old firms in our model economy face a lower passthrough from costs to prices due to the presence of the Kimball aggregator. When hit by a contractionary MP shock that decreases wages and puts a downward pressure on prices, dominant firms can cut prices by relatively less compared to young ones. Since markups depend on the ratio between firm prices and marginal costs, this mechanism is in turn responsible for the stronger countercyclical response in old firms' markups. In particular, we can match up to 20% of the empirically estimated relative difference in old and young firms' markups responses to a negative MP shock. In our analysis, we also show that the differential response of old firms in the model can be quantitatively decomposed to highlight the contribution of changes in aggregate variables to the overall general equilibrium impact on markups. In particular, the movements in real wages generated by a negative shock to the interest rate are found to be key in shaping the differential behavior of dominant firms' markups.

Finally, we conclude our quantitative analysis with an investigation of the shock amplification mechanisms at play in our framework, comparing our set up to a standard one-firm NK model with price rigidities. Both the presence of the Kimball aggregator and the heterogeneity of firms are shown to affect the way and extent to which MP shocks transmit in the economy, with output decreasing on average by roughly 20 percentage points (p.p.) more after a negative movement in the interest rate. Focusing on the role of the Kimball aggregator, since intermediate firms – especially old ones – temper their price drops after a negative MP shock due to the increase in their desired markup, the shock itself propagates more through quantities than through prices in our set up as opposed to the standard constant elasticity NK framework. At the same time, the Kimball aggregator alone is not sufficient to generate the observed amplification of MP shocks, as its effects on the elasticity of demand faced by firms kick in when firms are indeed heterogeneous and hence have a different passthrough from costs to prices. Firm heterogeneity is therefore key in affecting and amplifying the movements in the macroeconomic aggregates in the economy following a negative MP shock.

We see the contribution of this paper as twofold: on the one hand and to the best of our knowledge, we bring novel evidence on the remarkable heterogeneity in the response of firm markups to MP shocks based on corporate age. Specifically, while several empirical macroeconomic studies have focused on the different response of investment conditional on movements in the interest rates, we take a different perspective and explore the heterogeneous behavior of markups, the most direct measure of firms' market power. On the other hand, enriching a NK model with firm heterogeneity, demand accumulation and endogenous markups, we attempt to quantitatively study the role of firms' life-cycle in shaping the differential response of markups to changes in the interest rates, and then analyse how aggregate shocks propagate (and get amplified) in our model economy.

Related Literature. Our work builds on several macroeconomic contributions to the study of markups cyclicality. With respect to papers that have analysed the *aggregate* markup (see Gali et al. (2007), Hall (1988), Bils et al. (2018) and Nekarda and Ramey (2020)), we focus on the heterogeneous response of *firm-level* markups to changes in the interest rate, both from an empirical and quantitative point of view. Second, in comparison to recent research on firm-level markups by Hong (2017), Burstein et al. (2020), Meier and Reinelt (2020) and Alati (2020), we do no investigate markups response to business cycle movements, but rather markups behavior conditional on monetary policy.

On the other hand, we attempt to contribute to the theoretical and quantitative macroeconomic literature that has started incorporating micro-level heterogeneity into NK frameworks and understand its implications for the transmission of monetary policy. Recent studies in this field have focused on how household-level heterogeneity affects the consumption channel of monetary policy (see, for example, McKay et al. (2016), Kaplan et al. (2018), Auclert (2019), or Wong (2019)). More in line with the spirit of Ottonello and Winberry (2020)'s investigation of firm investment, we explore the role of firm-level heterogeneity in determining differences in the response of markups to monetary policy shocks. In so doing, we also relate our work to several analyses of supply-side heterogeneities in NK set ups, such as studies on price-setting behavior (see Golosov and Lucas (2007)), market power (see Klenow and Willis (2016) and Mongey (2017)), and product life-cycle (see Bilbiie et al. (2007) and Bilbiie et al. (2012)). With respect to these papers, we present a model of firm life-cycle behavior in order to examine the endogenous response of markups to monetary policy by firm age.

Our work is also related to Gilchrist et al. (2017), who study how financial distortions can create an incentive for firms to raise prices in response to adverse financial or demand shocks. While in Gilchrist et al. (2017) the rise in markups reflects firms' decisions to preserve internal liquidity and avoid accessing external finance, the endogenous response to markups in our set up is related to the differential demand elasticities faced by firms in their life-cycle. In this sense, we see our work as closely related to Baqaee et al. (2021) from a theoretical perspective: the authors explore the first-order effect on aggregate TFP caused by the reallocation of resources triggered by a demand shock across firms with non-uniform markups. While we answer a different research question, also in our model dominant firms tend to have both higher markups and lower pass-through from marginal costs to prices. When faced with an increase (decrease) in nominal marginal costs, high-markup firms raise (lower) their prices by less than low-markup firms in order to remain competitive.

Finally, our work is related to the macroeconomic literature pioneered by Gertler and Gilchrist (1994) that empirically documents how the effect of monetary policy can vary across firms of different characteristics. To the best of our knowledge, existing studies in this area have focused on firm-level investment, and assess how firm default risk (Ottonello and Winberry (2020)), liquidity (Jeenas (2019)) or age (Cloyne et al. (2018)) may shape the response of investment to monetary policy shocks. In a similar spirit, Fabiani et al. (2020) examine how monetary policy can influence the maturity structure of corporate debt. We also use state-of-the-art local projection techniques in our empirical analysis and explore the heterogeneous response of firms' markups to monetary policy shocks. In fact, our core contribution is to document that old firms' markups react more countercyclically to contractionary interest rate shocks than young ones, and then embed our finding into a heterogeneous firms NK framework augmented with endogenous markups formation.

The paper is organized as follows: in Section 2, we report and discuss the empirical evidence on the heterogeneous cyclicality of firms' markups after monetary policy shocks. Section 3 lays down our theoretical framework, characterized by heterogeneous firms in a NK setting with endogenous markups. Then, in Section 4, we illustrate the calibration and fit of the model, while in Section 5 we present steady state results and firm-level impulse responses to monetary policy shocks, and also discuss amplification mechanisms. In Section 6 we finally conclude and present the way ahead.

2 Empirical Analysis

In what follows, we study the heterogeneous cyclicality of firms' markups in response to monetary policy shocks. We begin by describing the sample of US firms and the monetary policy shock series on which we draw our evidence, and then illustrate how we estimate markups at the firm-level. Secondly, we document that old firms show a more countercyclical markups response after a monetary policy tightening. Finally, we briefly analyse the behavior of markups over firms' life-cycle and motivate why firm age could be a source of heterogeneity in markups responses to demand shocks.

2.1 Sample Construction

As previously mentioned, we make use of firm-level data from Compustat, which contains quarterly balance sheet information for North-American listed companies between 1975 and 2016. Compustat constitutes a panel of US corporations that is sufficiently high-frequency to be used to study monetary policy, and long enough to exploit within-firm variation. However, it comes at the expenses of representing the universe of publicly-listed incorporated firms only, even though these companies are estimated to make up for 30% of private sector employment. In terms of coverage, Compustat reports details on firm performance indicators and outcomes, including sales, liquid assets, financing sources, total assets, and production costs. It also reports the industry sector (SIC codes) where the business operates and firm age, which is the crucial dimension of heterogeneity in our analysis. Importantly, the age variable contained in the original dataset counts the years since incorporation, but we provide a robustness considering the establishment year for each firm in our sample.

Following standard practices in the literature, we restrict our attention to firms that are incorporated in the US and our final sample excludes utilities companies (SIC codes 4900-4999), financial entities (SIC codes 6000-6999), as well as corporations for which the industry code, or the information on sales, assets and production costs is missing. Whenever applicable, we deflate variables using a GDP-deflator from the NIPA tables. Table 1 reports summary statistics for the variables of interest.

Table 1: Summary Statistics

	Sales	Cogs	Assets	Leverage	Liquidity	Age
mean	447.69	303.17	4919.69	0.45	0.17	9.46
p25	6.06	3.31	37.83	0.04	0.02	4
p50	31.01	17.18	229.50	0.18	0.07	8
p75	164.58	100.60	1118.33	0.39	0.22	14
N	715,874	715,874	685,784	641,316	683,696	715,874

Notes: the first three columns are measured in millions of real 2012 $, while column (4) and (5) are ratios and column (6) is measured in years. *Cogs* is the cost of good sold, which includes production expenditures.

Our final sample of firms is then merged with two different interest rates datasets: first, we take the quarterly monetary policy shock series from Gürkaynak et al. (2005), who build a measure of interest rate surprises based on the % change in the FED Funds Futures rate in 30-minute windows around the policy announcement. Secondly, we also and primarily make use of the quarterly monetary policy shocks from Jarociński and Karadi (2020), a "pure" interest rate surprises series that removes from the estimation any component attributed to the provision of private FED information on the state of the economy to private agents through policy announcements. The common identifying assumption on the exogeneity in the variation of the policy rate is that nothing else occurs within this 30-minutes time window that could drive both private sector behavior and monetary policy decisions. Both series are available for the years and quarters between 1990Q1 and 2016Q4.

Before describing the estimation of markups at the firm level, we briefly recap on other important variables that we further employ as controls in our regressions. First, using balance sheet data, we compute the leverage and the holdings of liquid assets for the companies in our sample. With respect to the former, we take the ratio of corporate total debt divided by total assets in each period, both measured at book values and where debt is the sum of short term and long term debt. Parallel to that and to provide a measure of corporate liquidity, we compute the ratio of cash and short-term investments to total assets. Our main regression specifications also include firm size as a control, which is measured as the log of total assets (at book value).[1] Finally, we complement our firm-level data with general indicators of economic activity at quarterly level. In particular, we include the GDP growth rate, the Consumer Price Index (CPI) growth rate, the Excess Bond Premium (EBP), and the 1-Year Treasury rate change, all taken from the Federal Reserve of St.Louis (FRED) series.[2]

2.2 Markups Estimation

Firm-level markups are a common measure of whether companies are able to set their prices above marginal costs. To estimate them, we follow recent works by De Loecker and Warzynski (2012) and De Loecker et al. (2020), which are based on the production function approach pioneered by Hall (1988) on industry-level data. Their estimation strategy is grounded on firm's optimizing behavior with respect to production costs-minimization, and delivers an estimate of markups at the

[1] To eliminate seasonality, variables are measured as the rolling means in the previous 4 quarters as in Jeenas (2019).

[2] Even if the identified monetary policy shock series are exogenous, macro controls are included for robustness.

firm-level without specifying an explicit demand system. In fact, consider a firm i that employs a production technology given by:

$$Q_{i,t} = F_{i,t}(\boldsymbol{X}_{i,t}, K_{i,t}, \omega_{i,t})$$

where $\boldsymbol{X}$ is a vector of variable inputs, K is the predetermined input and ω is firm-specific productivity. The cost minimization problem for each producer can be hence expressed as follows:

$$\min_{\{\boldsymbol{X}_{i,t}, K_{i,t}\}} \{\boldsymbol{P}'_{i,t}\boldsymbol{X}_{i,t} + R_t K_{i,t} + \lambda_{i,t}(Q_{i,t} - Q(\cdot)\}$$

where $\boldsymbol{P}_{i,t}$ is the vector of prices for variable inputs, R_t is the price of the predetermined input, and $\lambda_{i,t}$ is the Lagrangian multiplier associated to the firm's cost minimization problem. One can then compute the first order condition (FOC) for a generic variable input $X^v \in \boldsymbol{X}$, which is given by:

$$\frac{\partial \mathcal{L}(\cdot)}{\partial X^v_{i,t}} = P^v_{i,t} - \lambda_{i,t} \frac{\partial Q(\cdot)}{\partial X^v_{i,t}} = 0 \tag{12}$$

Notice that the Lagrangian multiplier $\lambda_{i,t}$ can be also interpreted as the marginal cost of producing at a given level of output. Equation 12 can be further rearranged as:

$$\frac{\partial Q(\cdot)}{\partial X^v_{i,t}} \frac{X^v_{i,t}}{Q_{i,t}} = \frac{1}{\lambda_{i,t}} \frac{P^v_{i,t} X^v_{i,t}}{Q_{i,t}}$$

Defining the markup as price over marginal costs, $\mu_{i,t} \equiv \frac{P_{i,t}}{\lambda_{i,t}}$, it is possible to rearrange the FOC for a generic variable input $X^v \in \boldsymbol{X}$ such that it yields:

$$\mu_{i,t} = \theta^v_{s,t} \frac{P_{i,t} Q_{i,t}}{P^v_{i,t} X^v_{i,t}} \tag{13}$$

where $\theta^v_{s,t}$ is the elasticity of output with respect to the variable input X^v. The computation of markups can hence be implemented using firms' financial statements only. To estimate this theoretical expression in Compustat, we make use of both sales and cost of good sold data for each firm and in each quarter, which map to the denominator and numerator of Equation 13 according to:

$$\hat{\mu}_{i,t} = \widehat{\theta}^v_{s,t} \frac{\texttt{Sales}_{i,t}}{\texttt{Cogs}_{i,t}}$$

where we use the estimates of the sectoral output-input elasticity $\hat{\theta}^v_{s,t}$ from De Loecker et al. (2020).

2.3 Heterogeneous Markups Cyclicality

We then proceed to use Compustat quarterly balance-sheet data to investigate cross-sectional differences in the response of markups to interest rate policies. The main goal of our analysis is to estimate how firm i's markup $\mu_{i,t+h}$, at horizon $h \geq 0$, behaves in response to a monetary policy shock at time t, conditional on firm i's age just before the shock. To this end, we borrow the empirical strategies of Jeenas (2019) and Ottonello and Winberry (2020), and use a panel version of the Jordà (2005)'s local projections (hereafter: LP) to regress the cumulative difference in firm markups at different horizons on the interaction term between firm age at time $t-1$ and the monetary policy shock at time t, alongside a set of control variables. This flexible specification enables us to estimate impulse response functions on our firm-level panel data using the identified monetary shocks as instruments for changes in the policy interest rate. In particular, we estimate by OLS the following set of equations:

$$\Delta_h \log \mu_{i,t+h} = \sum_{x\in\mathcal{X}} \left(\alpha_{x,h} + \beta_{x,h}\Delta Y_{t-1} + \sum_{k=-\kappa}^{h} \gamma^k_{x,h}\varepsilon^m_{t+k} \right) \times \mathbb{1}_{i\in\mathcal{I}^x} \\ + \sum_{\ell=1}^{L} \boldsymbol{\delta}'_{\boldsymbol{h}} \boldsymbol{X}_{\boldsymbol{i,t-\ell}} + \varphi_{i,h} + \varphi_{s,t,h} + \boldsymbol{\vartheta}_{\boldsymbol{h}}\boldsymbol{t} + u_{i,t+h} \tag{14}$$

with horizons $h = 0, 1, \ldots, H$ and $H = 20$ quarters. The dependent variable is the cumulative change in markups for any firm i at horizon h, given by:

$$\Delta_h \log \mu_{i,t+h} \equiv \log \mu_{i,t+h} - \log \mu_{i,t-1}$$

Focusing on our regressors, $\mathbb{1}_{i\in\mathcal{I}^x}$ is an indicator that takes a value of 1 if $i \in \mathcal{I}^x$, namely if the firm i is above the *median* in one or more dimensions of the vector $x \in \{age, leverage, liquidity, assets\}$. The main coefficient of interest is $\gamma_{age,h}$, which captures the relative response of old companies (compared to young ones) to a variation in the FED short-term policy rate.[3] Note that we horse-race our main regressor – corporate age – against other layers of firm heterogeneity typically studied by the literature, such as size, leverage and liquidity (see Jeenas (2019)). We also interact the vector of firm-level regressors x with ΔY_{t-1}, which is the previous quarter's GDP growth, to control for the differential sensitivity of firm markups to the business cycle, following Ottonello and Winberry (2020).

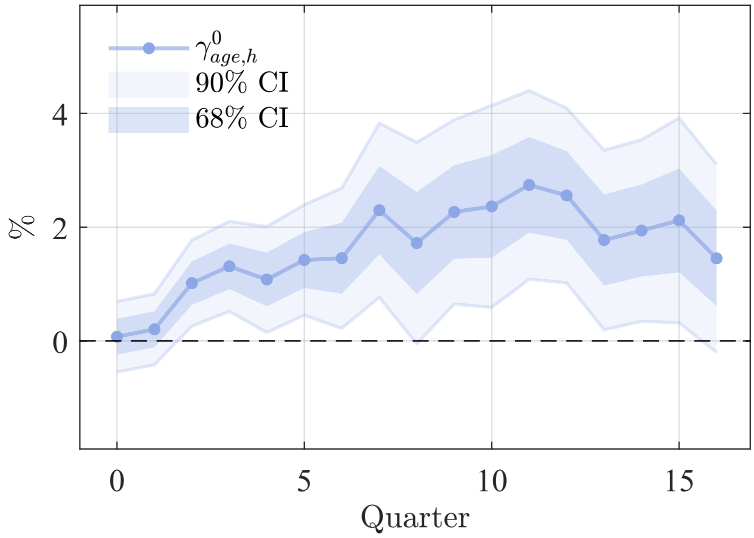
Furthermore, $\boldsymbol{\varepsilon}^{\boldsymbol{m}}_{\boldsymbol{t}} \equiv \sum_{k=-4}^{h} \varepsilon^m_{t+k}$ is the series of monetary policy shocks from Jarociński and Karadi (2020), while $\boldsymbol{X}_{\boldsymbol{i,t}}$ is a vector of controls that includes firm-level variables such as sales growth and overhead costs to sales, and macro-level controls like GDP and CPI growth, 1-year treasury rate change, EBP, and fiscal quarter dummies to account for seasonality. Following standard practices in the literature, we include control variable lags (up to 4) and measure the controls and the variables in vector x at the end of the quarter before the arrival of the monetary policy shock to ensure exogeneity with respect to it. We then allow for firm ($\varphi_{i,h}$), and sector-time ($\varphi_{s,t,h}$) fixed effects (FE) to control for the unobserved time-invariant heterogeneity at the level of the firm and to absorb time-varying shocks that are common to all firms in a given industry. We also include a linear and quadratic trend ($\boldsymbol{\vartheta}_{\boldsymbol{h}}\boldsymbol{t}$). Saturating the regression in Equation 14 with these FE implies that, first, our coefficients of interest are identified by within-firm variation over time, namely by changes in the markups response of an otherwise identical firm when it is old compared to when it was young. Secondly, our estimation fully exploits the cross-sectional variation across firms in a given industry. Finally, we cluster the standard errors at the firm and quarter level to account for correlation in the error term.[4]

As mentioned in the previous paragraph, our main coefficient of interest is given by $\gamma_{age,h}$, which captures the differential h-quarter growth of markups for firms above the median age after a 25 basis point hike in the interest rate (which corresponds to a rise of a quarter of a percent). Since we are including quarter FE and hence controlling for the time variation of the shock, the coefficient $\gamma_{age,h}$ can precisely identify the excess cyclicality of older firms' markups. In particular, Figure 2.1 reports the impulse response function obtained from the OLS-estimation of $\gamma_{age,h}$ in Equation 14,

[3]In our preferred specification, we therefore adopt a non-parametric estimation approach by using dummies instead of linear interactions. We show robustness checks following instead a parametric approach at the end of the section.

[4]Clustering at the firm level allows for a fully flexible dependence in the error terms across time within each company. Clustering by time is necessary whenever firm-level shocks are correlated within a quarter and if this effect may go potentially above the co-movement caused by industry-level shocks already captured by the sector-quarter dummies. We note that the confidence intervals on estimates would be significantly lower without clustering at the quarter-level.

Figure 2.1: Markups Response to a Monetary Policy Tightening



Notes: Within each quarter, firms' markups are winsorized at the 1% and 99% cutoff, to avoid any outlier to drive our results. Confidence intervals at 90% and 68%, which approximates one standard deviation.

along with standard confidence intervals around the point estimates. The magnitude of the $\gamma_{age,h}$ coefficient suggests that being above the median age before a contractionary monetary policy shock hits can imply up to a +3% statistically significant difference in the subsequent response of firm markups.

Interestingly, older firms' markups respond more countercyclically to a monetary policy tightening, with the cumulative effect lasting for at least 16 quarters after the shock. It is important to stress that we control for firm's FE and for other crucial determinants of between-firms heterogeneity studied by the literature, namely size, leverage and liquidity. Yet, none of the interactions between these three firm-level variables and the monetary policy shock are statistically significant predictors of markups heterogeneous response to interest rate changes, as further reported in the Appendix. Moreover, as in Cloyne et al. (2018), we note that firm age is pre-determined and cannot vary as a result of changes in monetary policy. In contrast, size, leverage and liquidity endogenously respond to shocks and vary over the business cycle, which can in turn affect the ranking of firms in the distribution of these variables. In this sense, even if there was any, it would be hard to interpret markups (ex-post) heterogeneity as being driven by ex-ante differences in these specific firm characteristics. Contrary to that, we establish that firm age can significantly determine the differential response of producers' markups to MP shocks, above and beyond other relevant firm characteristics.

The relative response of markups of old companies estimated through Equation 14 does not allow to understand the separate response of markups of firms in different age categories to monetary policy shocks. In particular, the regression specification in Equation 14 is saturated with industry and time FE that span out completely the time-series variation common across all firms. Hence, we proceed to estimate the following regression specification for firms above and below the median age:

$$\Delta_h \log \mu_{i,t+h} = \varphi_{i,h} + \boldsymbol{\vartheta_h t} + \sum_{\ell=1}^{L} \boldsymbol{\delta'_h X_{i,t-\ell}} + \sum_{k=-\kappa}^{h} \gamma^k_{x,h} \varepsilon^m_{t+k} + u_{i,t+h} \tag{15}$$

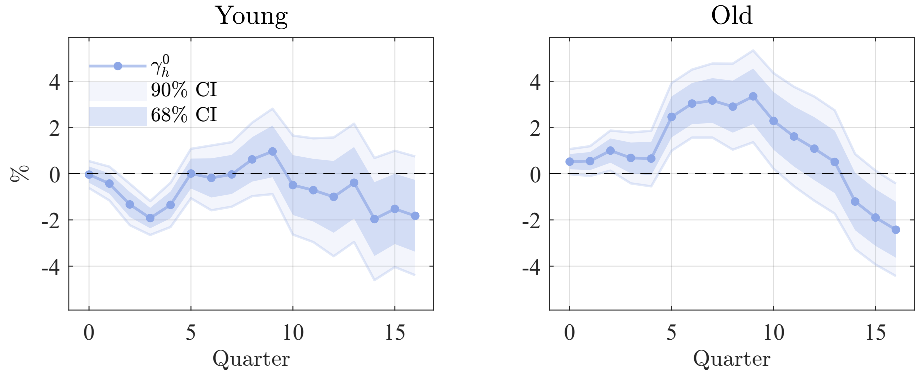
with horizons $h = 0, 1, \ldots, H$ and $H = 20$ quarters. Note that the dependent variable is the cumulative change in markups for any firm i at horizon h, which is defined as:

$$\Delta_h \log \mu_{i,t+h} \equiv \log \mu_{i,t+h} - \log \mu_{i,t-1}$$

Hence, in this second specification, we simply exploit the time-variation and look at the absolute change in markups after a change in the interest rate set by the FED for firms of different age categories, while the coefficient of interest γ_h is estimated for each age group separately. Note that $\boldsymbol{\varepsilon_t^m} \equiv \sum_{k=-4}^{h} \varepsilon_{t+k}^m$ is again the series of monetary policy shocks from Jarociński and Karadi (2020). Moreover, $\boldsymbol{X_{i,t}}$ is a vector of controls that include firm-level variables such as sales growth and overhead costs to sales, leverage, liquidity and assets, as well as macro-level controls like GDP growth, CPI growth, 1-year treasury rate change, EBP, and fiscal quarter dummies. Importantly, we also include control variable lags (up to 4). We allow for firm's FE ($\varphi_{i,h}$) to account for time-invariant firm-heterogeneity, and also add a linear and quadratic trend ($\boldsymbol{\vartheta_h t}$). Finally, we cluster our robust standard errors at the firm and quarter level to account for correlation in the error term.

The results of our estimation are shown in Figure 2.2: more specifically, the left panel documents the cumulative response of markups for firms below the median age to a negative movement in the FED interest rate, while the right panel focuses on the markups response of companies above the median age. This second estimation strategy further strengthens the insight from Figure 2.1, by documenting that older firms present a pronounced and statistically significant countercyclical response in their markups after a monetary policy tightening, while young firms' markups move procyclically, albeit the statistical significance of the estimated coefficient is much lower. Importantly, note that this second specification estimates a dynamic regression without the sector-time fixed effects and still shows that the above-median age firms' response in markups is nonetheless persistent, peaking 8 to 10 quarters after the shock. Taken together, these findings seem to suggest that dominant companies do adjust upwards their markups, whereas young firms' markups are generally less sensitive to monetary policy or tend to be adjusted downwards following a negative change in the FED interest rates. Finally, we check that our results hold more generally when we

Figure 2.2: Firms' Markups Response to a Monetary Policy Shock by Age Category



Notes: Within each quarter, firms' markups are winsorized at the 1% and 99% cutoff, to avoid any outlier to drive our results.

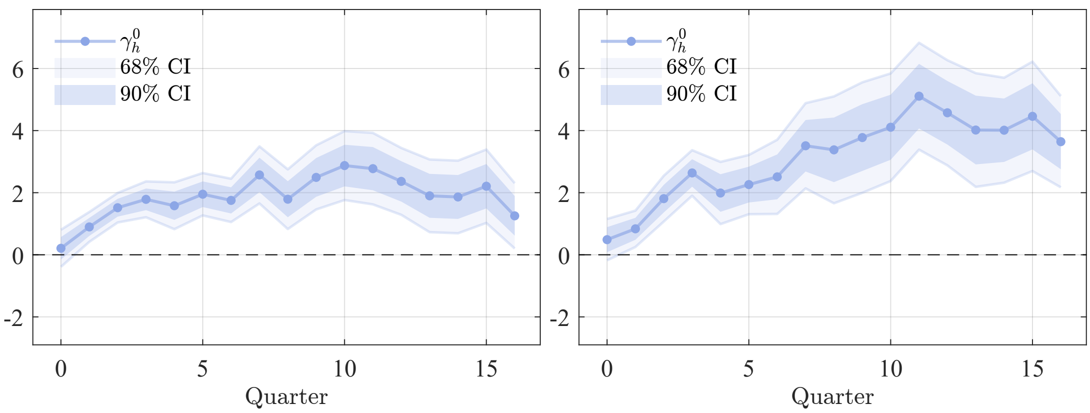
focus on a different partitioning of the age distribution, by for example considering as "old" those firms that are above the third quartile and as "young" firms all the others. Moreover, to have a further understanding of the possible interaction between corporate age and other firm-level characteristics, we also split old and young companies according to their position in the distribution of leverage, liquidity and size, which are other dimensions of firm's heterogeneity typically investigated in the literature that we have always controlled for in our regression analysis. As reported in Figure A.3, Figure A.4 and Figure A.5, corporate age is the crucial dimension determining the heterogeneous response of markups to monetary policy shocks, while other firm's characteristics – such as leverage, liquidity or size – are less powerful or even insignificant predictors of the differential behavior of markups at the company-level.

2.4 Discussion of Results

In what follows, we discuss the main robustness checks to further confirm the evidence on the role of age in shaping markups response to MP shocks at the firm-level. First, we run alternative regression specifications that present minor differences with respect to our baseline case. In particular, relative to the way we define the regressor of interest – namely corporate age – our main result is robust to consider age groups by industry and quarter, and also to interact the interest rate shock series with firm's age in a linear fashion, thereby adopting a parametric estimation strategy (similarly, we also linearly interact the MP shock series with the leverage, size and liquidity of the firms). The results of these alternative specifications are reported in Figure A.1 in the Appendix and both confirm that older firms present a stronger countercyclical response of markups to a monetary policy tightening.

Secondly, we check that our insights are not driven by the specific time span considered, in particular, by running again our estimation procedure on a sub-sample of the dataset that extends until the 2009 crisis. This is due to the fact that the Great Recession was indeed a period of exceptional financial conditions and, at the same time, the post-2009 era was characterized by a lower variation in the interest rate policy, with the federal funds rate often hitting the zero lower bound. However, as reported in the right panel of Figure 2.3, our results acquire a stronger statistical significance when the post-2009 era is excluded, and are hence not driven by specific period conditions only. Moreover, we also replicate our estimation using the monetary policy shock series

Figure 2.3: Using GSS Shocks (left) and Focusing on pre-2009 period (right)



Notes: Within each quarter, firms' markups are winsorized at the 1% and 99% cutoff, to avoid any outlier to drive our results.

from Gürkaynak et al. (2005) (hereafter GSS), which does not remove the informational component when measuring the interest rate surprises based on the 30-minute windows around FED policy announcements. As it is possible to check from the left panel of Figure 2.3, the coefficient on the interaction between the MP shock and firm's age – $\gamma_{age,h}$ – is economically relevant and significant, confirming that firms above the median age present an excess counter-cyclicality in their markups response to a monetary policy tightening, and that this differential effect lasts for an average horizon of 16 quarters after the shock.

Finally, our findings are robust to excluding future shocks from the estimation, as well as sector-quarter fixed effects, as reported in Figure A.2. In particular, future shocks were included among the regressors to control for the presence of auto-correlation and to increase the estimation precision, despite of the fact that the monetary policy shock series we have used should already be clear from confounding factors of this sort. Taking our results together, we argue that corporate age is a robust driver of firm's heterogeneity in the cyclicality of markups response to a monetary policy shock, and we hence proceed to briefly analyse the behavior of markups over firms' life-cycle.

2.5 Markups and Firm's Life-Cycle

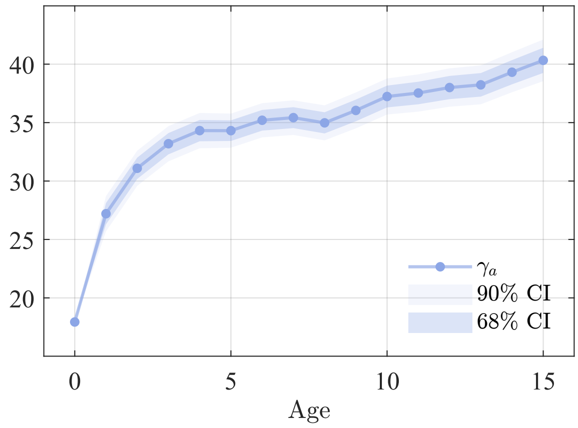
After having estimated the heterogeneous response of firms' markups to MP shocks, we provide a further discussion on why old and young companies may possibly show such stark differences in cyclical behavior of their respective markups. As mentioned before, corporate age has been investigated to be an important element of firm's employment and leverage dynamics over the business cycle by the works of Haltiwanger et al. (2013), Dinlersoz et al. (2018), and Pugsley et al. (2019). Interestingly, Cloyne et al. (2018) have studied how corporate age can determine investment heterogeneity across firms, especially in response to interest rate changes. Specifically, by documenting that the investment and the borrowing of younger firms paying no-dividends exhibit a large and significant decline in response to a tightening of the monetary policy, the authors argue that such companies are more likely to face financial frictions. In their view, this can also rationalize why the borrowing of young and non-dividend paying firms is far more sensitive to fluctuations in collateral values compared to other businesses, for which their results turn less significant.

In a similar in spirit, we argue instead that firm age matters significantly for the profile of markups and their response to a MP shock. In particular, it is reasonable to assume that corporate age may capture how *established* is a firm in her (unobservable) product market. Older firms, by means of having competed and produced in their given markets for a longer period, may be able to charge higher prices to consumers and hence be less subject to the downward pressure exerted on prices by a monetary policy tightening. In fact, according to the theoretical expression of firm-level markups, a negative interest rate shock puts a negative pressure on both input costs and prices. However, if older firms are able to decrease their prices by relatively less by taking advantage of their established position within a market, this may rationalize a more countercyclical markups response to a monetary tightening. To provide suggestive evidence of how corporate age is related to firm's established position in a given market, we examine the profile of markups ($\mu_{i,t}$) and selling expenditure ($sr_{i,t}$) over the business life-cycle using Compustat data. In particular, we run the following regressions:

$$\log \mu_{i,t} = \alpha + \sum_{a=2}^{\mathcal{A}} \gamma_a \mathbb{1}_{\{age_{i,t}=a\}} + \varphi_{s,t} + \varepsilon_{i,t}$$

where $\varphi_{s,t}$ are sector and quarter fixed effects. Not only old firms are on average big, but they most importantly tend to have higher markups, as documented in Figure 2.4. The main takeaway from Figure 2.4 is that firms are able to charge higher markups (hence higher prices) as they grow older. We interpret this suggestive evidence as an indicator that older companies may have already secured their customer base enough to be less inclined to drastically reduce prices in response to a negative monetary policy shock, resulting in the stronger markups counter-cyclicality documented in the previous paragraphs. Since we argue that firm age is a proxy for how established a production unit is in her given market, we will then rationalize our findings into a NK model with heterogeneous firms, demand accumulation and endogenous markups that not only will generally differ across producers but that will be further allowed to grow according to the firm's life-cycle.

Figure 2.4: Markups over Firm's Life-Cycle



3 The Model

In this section, we outline our theoretical framework and discuss how each assumption relates to and can deliver qualitative predictions in line with the evidence from the data. In particular, we enrich a relatively standard NK to accommodate three main novelties: first, we allow for full heterogeneity on the supply side of the economy, by including heterogeneous intermediate firms that produce a different variety of input used in the final good sector. Secondly, we introduce a simple form of demand accumulation that makes the demand for the good of a given firm increase with the time that the firm survives on the market. Third, we embed endogenous and variable markups in the economy, which differ across companies according to the quantities they produce, and that also evolve along with the life-cycle of the firms. We now proceed to present the model in full details below.

3.1 Household's Side

Time is continuous. The model features a representative household that optimizes the discounted flow of utility from consumption and labor over an infinite lifetime horizon, where we indicate the discount factor as $\rho \geq 0$. We assume that the utility of the agent is strictly increasing and concave in consumption, and strictly decreasing and convex in the amount of hours worked respectively. Preferences are time-separable and the infinite stream of household's utility is hence given by:

$$\int_0^\infty e^{-\rho t}\left(\frac{C_t^{1-\nu}}{1-\nu} - \frac{L_t^{1+\gamma}}{1+\gamma}\right)$$

where ν represents the risk aversion in the CRRA utility function over consumption, whereas γ is the inverse of the Frisch labor elasticity. Moreover, $L_t \in [0,1]$ are the hours supplied as a fraction of the time endowment (normalized to 1), while C_t denotes the aggregate consumption good. In each period, the household can borrow in bonds B_t at the real interest rate r_t. Finally, the household owns all the firms in the economy, while labor supply, aggregate consumption and bond investment paths are chosen as a result of a value maximization problem subject to a standard budget constraint:

$$\mathcal{V} = \max_{\{C_t, L_t, \dot{B}_t\}} \int_0^\infty e^{-\rho t} \left(\frac{C_t^{1-\nu}}{1-\nu} - \frac{L_t^{1+\gamma}}{1+\gamma} \right) dt$$
$$\text{s.t.} \quad C_t + \dot{B}_t = W_t L_t + r_t B_t + D_t$$

where we denote by D_t the dividends from the firms and by W_t the wage earned by the household in real terms. As we will explain below, r_t will be determined by the monetary policy and Fisher equation, while W_t is determined by the market clearing condition for labor. Solving for the optimal value of consumption and labor, we get the following standard Euler and labor supply equations:

$$r = \rho + \nu \frac{\dot{C}}{C} \tag{16}$$

$$L^\gamma C^\nu = \frac{W}{\varphi} \tag{17}$$

3.2 Final Good Producer

A competitive representative final-good producer aggregates a continuum of intermediate inputs indexed by $i \in [0, 1]$ according to the following expression:

$$\int_0^1 \mathcal{K}\left(a_{i,t} \frac{y_{i,t}}{Y_t} \right) di = 1 \tag{18}$$

where we assume that intermediate inputs denote by y_t are aggregated using the Kimball aggregator $\mathcal{K}$, with $\mathcal{K}'(\cdot) > 0$, $\mathcal{K}''(\cdot) < 0$, and $\mathcal{K}(1) = 1$. Notice that the CES aggregator obtains as a special case of the Kimball aggregator, and namely when $\mathcal{K}(q) = q^{\frac{\sigma-1}{\sigma}}$ for an elasticity of substitution $\sigma > 1$. Importantly, $a_{i,t}$ is a stochastic demand process that will be explained in due details in the next paragraph. For the moment, taking the prices $p_{i,t}$ of any intermediate input i as given and normalizing the price of the final good to 1, the final good producer minimizes production costs subject to Equation 18. The optimality condition of this problem gives rise to the *inverse demand* function for good i:

$$p_{i,t} = \mathcal{K}'\left(a_{i,t} \frac{y_{i,t}}{Y_t} \right) a_{i,t} \mathcal{D}_t \tag{19}$$

where:

$$\mathcal{D}_t = \left(\int_0^1 \mathcal{K}'\left(a_{i,t} \frac{y_{i,t}}{Y_t} \right) a_{i,t} \frac{y_{i,t}}{Y_t} di \right)^{-1} \tag{20}$$

is a *demand index*. In the CES case $\mathcal{K}(q) = q^{\frac{\sigma-1}{\sigma}}$ this index is a constant, so that $\mathcal{D}_t = \frac{\sigma}{\sigma-1}$ and Equation 19 reduces to the familiar constant elasticity demand curve given by $p_{i,t} = \left(a_{i,t} \frac{y_{i,t}}{Y_t} \right)^{\frac{-1}{\sigma}}$. To keep the exposition concise, further derivations related to Equation 19 and Equation 20 are contained in the Appendix. Moreover, we use the Klenow and Willis (2016) specification for

$\mathcal{K}(q)$ given by:

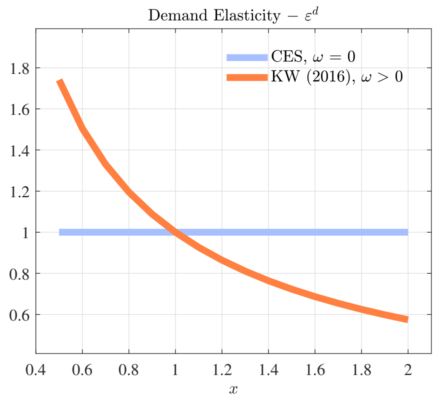
$$\mathcal{K}(q) = 1 + (\sigma - 1)\exp\left(\frac{1}{\omega}\right)\omega^{\frac{\sigma}{\omega}-1}\left[\Gamma\left(\frac{\sigma}{\omega}, \frac{1}{\omega}\right) - \Gamma\left(\frac{\sigma}{\omega}, \frac{q^{\omega/\sigma}}{\omega}\right)\right] \tag{21}$$

where $\sigma > 1$, $\omega \geq 0$ and $\Gamma(s,x)$ is the upper incomplete Gamma function such that $\Gamma(s,x) := \int_x^\infty t^{s-1}e^{-t}dt$. In particular, ω is the *super elasticity*, which is 0 in the CES aggregator. Finally, we can derive an analytical expression for the elasticity of demand ε_i^d that is faced by a producer of any good variety i as a function of the relative quantity of good i in the economy, which is given by:

$$\varepsilon_i^d = \sigma\left(a_{i,t}\frac{y_{i,t}}{Y_t}\right)^{-\frac{\omega}{\sigma}}, \quad \omega \geq 0$$

As already pointed out, the standard CES case is recovered when $\omega = 0$ and hence the elasticity of demand $\varepsilon_i^d = \sigma$ is constant across producers. In contrast, in the case of Kimball aggregator the elasticity of substitution is lower for firms with higher relative quantity $x = a\frac{y}{Y}$, implying that larger firms can choose higher markups, in a similar spirit to the different set up adopted in Atkeson and Burstein (2008) and as further made clear in Figure 3.1. When $\omega > 0$, the extent to which a firm's markup increases with its relative size is determined by the ratio σ/ω, which will also be shown to quantitatively matter in shaping how markups change with monetary policy later on in the analysis.

Figure 3.1: Kimball Aggregator



3.3 Intermediate Good Producers

Each intermediate good i is produced by a monopolistically competitive firm using effective units of labor $\ell_{i,t}$ in the production process and according to the technology given by:

$$y_{i,t} = \ell_{i,t}^{1-\alpha} \tag{22}$$

with $\alpha \in [0,1]$. In each time t, firms hire labor at wage W_t in a competitive labor market. As already mentioned, intermediate producers are monopolistic competitors on their respective markets and each one of them faces a demand function which can be written explicitly from Equation 19 as:

$$y_{i,t} = \left(1 - \omega\log\left(\frac{\sigma}{\sigma-1}\frac{1}{a_{i,t}}\frac{p_{i,t}}{\mathcal{D}_t}\right)\right)^{\sigma/\omega}\frac{Y_t}{a_{i,t}} \tag{23}$$

Moreover, each intermediate firm i is characterized by a process of demand accumulation given by a_i, which shows some persistence ρ_a and an idiosyncratic risk component given by $\xi_a d\mathcal{W}$ (as we work in continuous time, note that $d\mathcal{W}$ is a standard Wiener process). We also include a drift $\bar{a}$ that allows for the demand to grow over time, generating a realistic life-cycle profile. It is important to stress that we load the heterogeneity across firms that we see in the data in this specific process, which is meant to capture in a reduced-form the fact that markups and size increase with the firm's life-cycle. Such demand process may actually rationalize some underlying form of customer accumulation or, alternatively, a latent phenomenon of consumers habit formation. In other words, one can think about it in the sense that the more consumers experience the good of a given firm i, the more inelastic their demand for that specific item would consequently be.

Intermediate firms in this economy, characterized by the demand process a, maximize the discounted stream of profits by choosing prices. Hence, at each instant in time, the state of the economy is given by the joint distribution $\lambda_t(da, dp)$. Finally, intermediate producers discount future profits at the rate $r_t + \delta$, where δ is the exogenous Poisson intensity that determines firm's exit. Exiters are replaced by new firms with an initial a_0 drawn from a log-normal distribution of mean $\bar{a}_{entry}$ and standard deviation $\xi_{a,entry}$, which will be further discussed in the calibration exercise. Moreover, intermediate firms bear Rotemberg adjustment costs when changing prices, which we assume to be proportional to their sales and quadratic. We can summarize the problem of a given firm i as follows:

$$
\begin{aligned}
\mathcal{J}_{i,0} = & \max_{\{\dot{p}_{i,t}, \ell_{i,t}, y_{i,t}\}_{t\geq 0}} \mathbb{E}_0 \int_0^\infty e^{-\int_t^\infty (r_t+\delta)dt} \left\{ p_{i,t}y_{i,t} - W_t \ell_{i,t} - \frac{\vartheta}{2}\left(\pi_t + \frac{\dot{p}_{i,t}}{p_{i,t}}\right)^2 p_{i,t}y_{i,t} \right\} dt \\
\text{s.t.} \quad & y_{i,t} = \left(1 - \omega \log\left(\frac{\sigma}{\sigma-1}\frac{1}{a_{i,t}}\frac{p_{i,t}}{\mathcal{D}_t}\right)\right)^{\sigma/\omega} \frac{Y_t}{a_{i,t}} \\
& y_{i,t} = \ell_{i,t}^{1-\alpha} \\
& \dot{a}_{i,t} = \rho_a(\bar{a} - a_{i,t})dt + \xi_a d\mathcal{W}_{i,t} \\
& p_{i,0} \text{ and } a_{i,0} \text{ given}
\end{aligned}
$$

Importantly, the initial price set by entrant firms p_0 is the one that maximizes the expected value $\mathcal{J}_{i,0}$ for a given initial value of firm's productivity $a_{i,0}$. Note that, in the solution process, the demand process given by $\dot{a}_{i,t}$ is exponentiated. Intermediate firms take as given equilibrium paths for the real wage $\{W_t\}_{t\geq 0}$ and the interest rate $\{r_t\}_{t\geq 0}$. In steady state, the recursive solution to this problem consists of decision rules for labor $\ell(a, p; \mathcal{S})$ and output $y(a, p; \mathcal{S})$, with $\mathcal{S} := (r, W, Y, \mathcal{D}, \pi)$. These rules in turn also imply optimal drifts for prices, and together with a stochastic process for a, induce a stationary joint distribution of firms given by $\lambda(da, dp; \mathcal{S})$ and characterized by a standard Kolmogorov forward equation. Out of the steady state, each of these objects is time-varying and depends on the time path of prices and policies: $\{\mathcal{S}_t\}_{t\geq 0} := \{r_t, W_t, Y_t, \mathcal{D}_t, \pi_t\}_{t\geq 0}$.

3.4 Monetary Authority

Our model economy features a monetary authority that sets the nominal interest rate according to a standard Taylor rule, penalizing deviations from the optimal inflation rate π^* in the following way:

$$
i_t = \phi_\pi(\pi_t - \pi^*) + \rho + \varepsilon_t^m
$$

where $\phi_\pi > 1$, ρ is the discount factor and ε_t^m is the monetary policy shock that can be mapped directly to the series from either Jarociński and Karadi (2020) or Gürkaynak et al. (2005) that we have used in the empirical analysis of the paper. Note that $\varepsilon_t^m = 0$ in steady state: one of our main quantitative exercises will be precisely to study the economy's adjustment after an unexpected temporary monetary shock, namely after a change in ε_t^m. Finally, given inflation π_t and the nominal interest rate i_t, the real return on bonds r_t is determined by the Fisher equation $r_t = i_t - \pi_t$.

3.5 Equilibrium Condition

An equilibrium in this economy is defined as a set of paths for individual household's $\{C_t, L_t\}_{t\geq 0}$ and firm's decisions $\{\dot{p}_{i,t}, \ell_{i,t}, y_{i,t}\}_{t\geq 0}$, input prices $\{W_t\}_{t\geq 0}$, the return on bonds $\{r_t\}_{t\geq 0}$, the inflation rate $\{\pi_t\}_{t\geq 0}$, the distribution of firms $\{\lambda_t\}_{t\geq 0}$, the demand index $\{\mathcal{D}_t\}_{t\geq 0}$, and aggregate quantities such that, at every t: (i) the household and the firms maximize their objective functions taking as given equilibrium prices and aggregate quantities; (ii) the sequence of distributions satisfies aggregate consistency conditions; (iii) all markets clear. There are three markets in our economy: the bond market, the labor market, and the goods market. The bond market clears when the following holds:

$$B_t = 0 \tag{24}$$

Moreover, the labor market clears when:

$$L_t = \int \ell_t(a,p)d\lambda_t$$

Finally, the goods market clears according to:

$$C_t = Y_t - \int \frac{\vartheta}{2}\left(\pi_t + \frac{\dot{p}_t(a,p)}{p}\right)^2 py(a,p)d\lambda_t$$

where C_t is the total real expenditure in consumption, Y_t is aggregate output, and the last term is the sum of adjustment costs to prices paid by intermediate firms.

4 Quantification

In what follows, we proceed to explain the quantification of our model, including the calibration strategy and the overall fit of both targeted and untargeted moments computed from available US data. In particular, we discuss the ability of our theoretical framework to replicate salient features of the markups and firms' distribution, which is a crucial property needed to provide a link with the empirical analysis of the previous sections. Once quantified, the model is then used in Section 5 to study and analytically decompose the impulse response functions of firms' markups after a negative monetary policy shock. Moreover, in Section 5, we also compare the amplification mechanism implied in our framework with respect to a standard representative firm New Keynesian model.

4.1 Calibration

A model period in one quarter. Of the 14 parameters we need to calibrate, 8 are fixed outside of the model, for which we pick common values used in the literature. In particular, we set the risk aversion $\nu = 2$ and the disutility of labor $\gamma = 2$, while the discount factor $\rho = 0.012$ is specified to deliver a yearly interest rate of 5% in equilibrium. With respect to the parameters related to firms' life-cycle, technology and pricing behavior, we fix the quarter exit rate $\delta = 0.024$ to imply that 10% of the firms exit each year, and the returns to scale $\alpha = 0.33$ such that the labor share is around 0.6 in equilibrium. Moreover, it is important to specify that we normalize at 1 the mean demand $\bar{a}_{entry}$ faced by entrant intermediate firms, while the demand dispersion $\xi_{a,entry}$ at entry is set to be equal to the dispersion of the demand process faced by incumbents.[5] Finally, the monetary policy coefficient $\phi_\pi = 1.5$ in the Taylor rule is chosen to replicate a similar strategy as in Taylor (1999) and Galí (2015). The full list of both fixed and fitted parameter, as well as targeted moments, is presented in Table 2.

Table 2: Estimated Parameters and Targeted Moments

Fixed	Value	Description	Moments	Model	Data
ρ	0.012	Discount factor			
ν	1	Risk aversion			
γ	2	Inverse Frisch elasticity			
α	0.33	Production function curvature			
δ	0.024	Exit rate			
$\bar{a}_{entry}$	1	Mean demand entrants			
$\xi_{a,entry}$	0.11	Demand dispersion entrants			
ϕ_π	1.5	Taylor rule coefficient			
Fitted	**Value**	**Description**	**Moments**	**Model**	**Data**
θ	20	Price adjustment cost	Avg. cost change prices over sales	0.11	0.09
σ	4	Elasticity of demand	Avg. markup	1.68	1.68
ω	5.1	Superelasticity of demand	Elasticity markups to sale shares	0.11	0.10
$\bar{a}$	2	Mean demand	Median markup	1.37	1.30
ξ_a	0.11	Demand dispersion	Markups standard deviation	1.23	1.22
ρ_a	0.02	Demand mean reversion	Markups growth between age 0-5	0.24	0.22

Note: Estimates for fitted parameters from Compustat Data (1990Q1-2016Q4). For the fixed parameters, see text.

In addition to that, we need to endogenously assign values to the remaining 6 parameters, for which we match as many salient moments from available US data. To begin with, we set the price adjustment cost factor $\theta = 20$ such that the average ratio between the cost paid by firms to change prices and their sales is the same in the model and in the data.[6] As standard in the literature, we set the elasticity of demand $\sigma = 4$ to match an average markup of 1.68 computed in

[5] Our results do not depend on this choice, which is just a simplification for the sake of the estimation procedure.

[6] Estimates vary between 0.04 for physical costs and 0.09 for customer costs, see for example Levy et al. (1997) and Zbaracki et al. (2004). As in Golosov and Lucas (2007) and Baley and Blanco (2019), we choose a value in between.

the sample of Compustat firms:[7] this parameter determines the level of substitutability across the output of different producers in the model, and hence influences the average market power in the economy. Moreover, the superelasticity of demand ω is fitted such that the elasticity of markups to sale shares in the model is the same as in the data. In particular, our choice is motivated by the fact that the parameter ω in the Kimball aggregator is tightly linked to the relationship between the relative size of the firms and their markups: if ω was to be 0, such relationship would be null because all firms would have the same markup independently of their size. On the contrary, for $\omega > 0$, the higher the ω the higher the dependence of markups on sales shares. To this end, using Compustat firm-level data, we empirically estimate the elasticity of (log) markups to (log) sales shares according to:

$$\log \mu_{i,t} = \beta * \log(sales\ shares)_{i,t} + \varphi_{s,t} + \varepsilon_{i,t} \tag{25}$$

where $\varphi_{s,t}$ are sector-time FE and the coefficient β precisely informs by how much markups are linked to firms' sales shares. In the model, we use the theoretical definitions of markups and sales shares.

Finally, turning to the parameters related to the demand accumulation process, the mean demand is set to match the median markup in the US economy, as $\overline{a}$ identifies the distance between the average demand faced by entrants and incumbents, and hence relates to the skewness of the markup distribution. Furthermore, the dispersion in the demand process faced by incumbent firms ξ_a is identified from the standard deviation of markups, while the mean reversion in the demand process ρ_a is picked to match the growth of markups for firms between age 0 to 5. In particular, a higher mean reversion in the demand process impacts how fast firms grow, and therefore relates to the trajectory of markups over the firm's life-cycle.

4.2 Quantitative Fit

In the following paragraphs, we present and discuss our main validation exercises, which provide a overview of the quantitative fit of our framework with respect to empirical moments and data features that have not been targeted in the calibration. In particular, we first discuss the cross-sectional and life-cycle characteristics of firms in our model, and how they compare to their empirical counterparts from Compustat. Secondly, we dig into the properties of the markup distribution and then analyse markups dynamics over the firm's life-cycle. Finally, we conclude with a note on the model and data-implied elasticity of wages to sales and relate it to the behavior of markups under the Kimball aggregator case and in imperfect competition, following similar lines as in Edmond et al. (2018).

4.2.1 Implications for Markups in Steady State

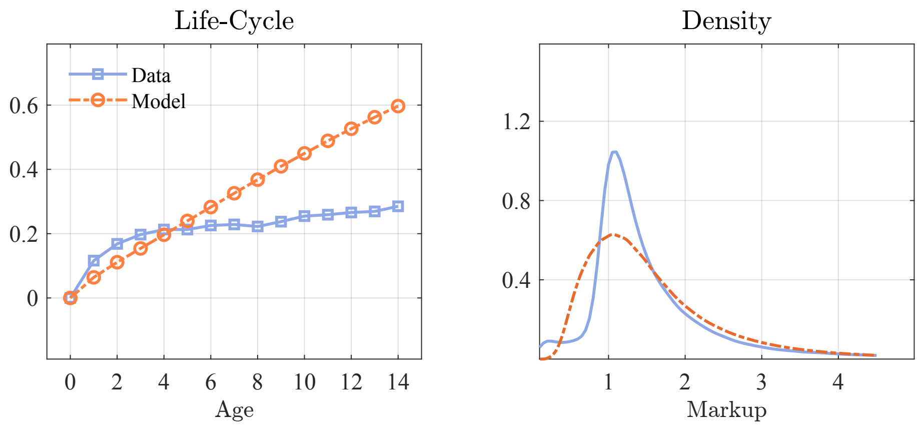
One of our main validation exercises is to look at the properties of markups in the data and compare them with the ones implied by our quantitative framework. Importantly, in Section 2, we have shown that markups increase with the age of the firm, and argued that such behavior may in principle be due to the fact that, as businesses advance along their life-cycle, they are also able to establish their position in their respective markets and progressively accumulate demand for their products. This in turn allows producers to progressively charge higher prices and hence set higher markups. Accordingly, the left panel in Figure 4.1 reports the pattern of markups over firms' life-

[7] We use Compustat Data between 1990Q1 and 2016Q4. For the empirical definition of markups, see Section 2.

cycle both in the model and in the data. In particular, we remind the reader that the empirical series has been computed using Compustat data between 1990Q1 and 2016Q4 and netting out sector and time FE.

On the one hand, the model slightly underestimates the rapid increase of markups in the first 5 years of a firm's life, whereas it tends to modestly overestimate their subsequent growth in the next years.[8] On the other hand, our calibrated framework can replicate qualitatively the growth of markups over firm age and match more than half of the quantitative features of the relationship between markups and the life-cycle of producers. Importantly, it needs to be stressed that the ability of the model to imply life-cycle markups' properties consistent with the empirical observations will prove crucial when assessing the differential response of firms to interest rate shocks. In fact, as documented in Section 2, old firms' markups show a more countercyclical response after a negative MP shock: absent the fit of the life-cycle profile of markups, our model would then not be able to replicate the heterogeneous response of markups to a MP shock according to firms' relative age.

Figure 4.1: Markups Steady State Properties



Secondly, as illustrated in the right panel of Figure 4.1, we can reasonably match the entire distribution of markups estimated from Compustat data. In particular, while a couple of distributional properties have been indeed targeted in the calibration, the model itself delivers a fat right tail in the distribution of markups consistent with our empirical observations and with the analysis of De Loecker et al. (2020). As reported in Table 3, our quantitative framework implies that the bottom 25% firms in the distribution have an average markup of 1.15, against an empirical value of 1.03 computed in the data, while a similar fit holds for the top 75% firms. Matching the right proportions of high and low-markup firms' will prove crucial when comparing the response of firms' markups to a monetary policy shock across companies that are below or above the median age.

4.2.2 The Link between Wages and Sales

While the superelasticity parameter ω has been identified by computing the elasticity of markups to sales shares, our calibrated model has also a testable prediction on the relationship between the wage bill and the sales of firms, which we can match as an untargeted dimension. In particular,

[8]The fit is very precise during the first years of business operations which is due to the fact that, in our calibration, we target the mean reversion in the demand process ρ_a to match the growth of markups for firms aged 0 to 5.

Table 3: Distributional Properties of Markups

	Model	Data
Bottom 25% Firms	1.15	1.03
Top 75% Firms	1.79	1.86

recall that markups are a measure of whether firms can set prices above their marginal costs. Similarly to Edmond et al. (2018), in our theoretical set up the salaries paid by firm i hence depend on its sales and markup according to a simple expression given by:

$$\text{wage bill} = \frac{\text{sales}}{\text{markup}}$$

Moreover, if the superelasticity ω in the Kimball aggregator was equal to zero as in the standard NK model, markups would not increase with firm sales and, in turn, the wage bill shares would increase one-for-one with sales shares. But when ω is strictly positive, as in our framework, markups do increase with firms sales, implying that the wage bill increases less than one-for-one with sales. In this sense, both empirically and quantitatively, the extent to which the wage bill share of firms increases with their sales shares can therefore be linked to the extent to which markups increase with producers' size. A small caveat to keep in mind is that Compustat does not report a precise measure for firms' wage bills but only a balance sheet item related to the cost of goods sold. This variable comprises the cost of all variable inputs used in production, included (but not exclusively) labor. Nevertheless, we exploit the available data to run the following regression:

$$\log(\textit{wage bill shares})_{i,t} = \beta * \log(\textit{sales shares})_{i,t} + \varphi_{s,t} + \varepsilon_{i,t} \tag{26}$$

where $\varphi_{s,t}$ are sector-time FE and the coefficient β precisely informs by how much variable input costs are linked to firms' sales shares. A value of the elasticity $\beta < 1$ confirms the fact that, absent perfect competition – as in our model –, firms increase sales by increasing prices, thereby suppressing produced quantities. In turn, this mechanism implies that growing firms also demand less employment, which creates a wedge such that wage bill shares do not move one to one with sales shares. The results of the empirical estimation and quantitative fit are reported in Table 4.

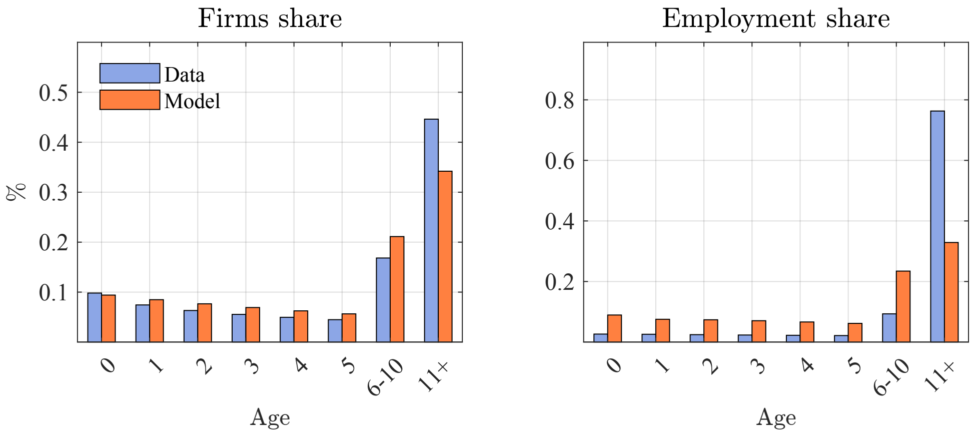
Table 4: Estimated Relationship between Wages and Sales

	Model	Data
Elasticity of Wage Bill Shares to Sales Shares	0.87	0.88

4.2.3 Cross-sectional and Life-Cycle Properties

In our last exercise, we analyse the distribution of firms by age and the life-cycle profile of both employment and sales growth rates for the businesses in our model economy. In Figure 4.2, we report the distribution of firms and employment shares by age, comparing the empirical ones from Compustat (1990Q1-2016Q4) to the ones obtained in our quantified framework. Note that none of these distributions was targeted in the calibration of the model, and hence both comparisons are to be considered as a pure validation exercise. First, focusing on the left panel, one can observe that our framework succeeds in replicating the distribution of firms over their age, and only partially underestimates the share of businesses that are 11+ years old. In this sense, as most of our empirical analysis is highly focused on markups' properties over the life-cycle of firms, it is remarkably important that we are able to capture the correct number of firms per age bin. In fact, the share of companies in each age bin influences the heterogeneous response of markups' to monetary policy shocks, and hence is relevant to get a correct quantitative fit of the empirically estimated dynamics of markups by firms' age.

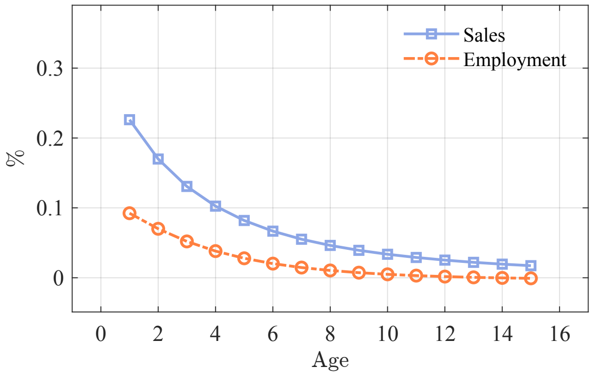
Figure 4.2: Distributions of Firms and Employment Shares by Age



Secondly, in the right panel of Figure 4.2, we plot the distribution of employment shares over firm age, comparing the empirical ones with their model-implied counterparts. As it becomes clear from the graph, our framework is able to match only up to half of the right tail in the employment share distribution. This is precisely due to the fact that, in the model economy, big firms (and hence old firms), find optimal to increase sales by increasing prices, thereby suppressing produced quantities and employment demand. This mechanism is a key characteristic of our set up in which companies operate in an environment with imperfect competition, and it is hence responsible for the fact that 11+ years old firms in the model generate a lower employment share compared to their empirical counterpart. Nonetheless, from this particular validation exercise we are still able to get a satisfactory fit of both firms and employment share distributions over the age of the businesses.

As a final note, in Figure 4.3 we plot the average employment and sales growth rates over the life-cycle of firms. Understandably, both measures decrease over time, as companies become old and hence slow down in their growth processes: this means that growth rates are unconditionally negatively correlated with age, as empirically noted in Dunne et al. (1989). However, sales grow relatively more than employment, which is indeed consistent with the early discussion related to the employment share distribution depicted in the right panel of Figure 4.2. In particular, as argued in the previous paragraphs and due to the presence of the Kimball aggregator, markups do increase with firms sales, implying that the wage bill increases less than one-for-one with sales, depressing the labor demand by firms and resulting in lower employment growth rates compared to the growth rate of firm's sales. In other words, due to market power, companies can increase sales by raising prices and decreasing output, which lowers their demand of labor and hence employment growth.

Figure 4.3: Employment and Sales Growth Rates



5 Results

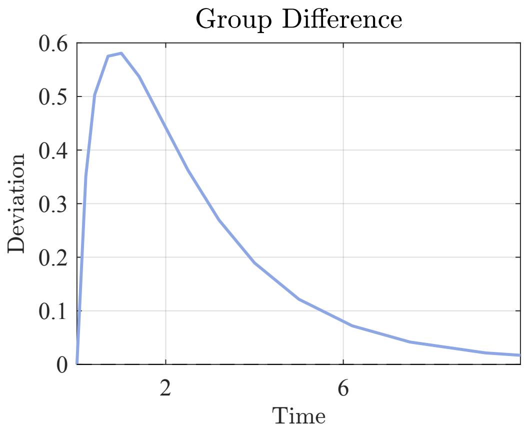
In the following section, we begin by discussing the response of firm markups to interest rate shocks, and compare the relative response of old and young firms in the model with the ones obtained in the data and reported in Section 2. Secondly, having assessed how much of the heterogeneity in response of markups to interest rates by firm age our model is able to replicate, we also illustrate by how much the changes of aggregate variables such as output and wages after a MP shock contribute to the differential response of markups of old firms with respect to young ones. Finally, we conclude by analysing the amplification of shocks at work in our framework compared to a one-firm NK model.

5.1 Response of Markups to Monetary Policy Shocks

We proceed to illustrate the dynamics of the economy after the arrival of a negative MP shock and to compare the relative response of old firms and young firms' markups to the ones obtained from the data and discussed in Section 3. As standard in frameworks characterized by nominal rigidities, a negative MP shock features an increase in the nominal interest rate and implies a downwards pressure on the labor cost W. Parallel to that, both employment, consumption and output decrease on impact and slowly recover as the shock fades away, while the downwards pressure on prices determines a deflationary episode. Moreover, the aggregate markup increases as a result of decreasing labor costs, and hence shows a countercyclical behavior in response to negative shocks to the nominal interest rate. The aggregate response of our calibrated economy hence resembles qualitatively the one of a standard NK textbook model, as in Galí (2015). However, the aggregate pattern of markups masks a noticeable degree of heterogeneity at the firm-level which we explore in what follows.

To obtain a comparable set up to our empirical analysis, we first categorize firms in our model economy by their age decile and then classify all businesses above the median age as "old" and below the median age as "young". We simulate the hit of a negative MP shock, otherwise defined as an exogenous increase in the nominal interest rate. Similarly to the empirical analysis in Figure 2.1, we then compute the differential response of markups to a MP shock for firms above the median age compared to firms below the median age. Figure 5.1 plots the differential response of markups by firm age over a horizon of several quarters and in deviation from the mean response. Clearly, firms above the median age respond more countercyclically to a negative MP shock compared to businesses below the median age, consistent with the empirical evidence documented in Compustat data.

Figure 5.1: Markups IRFs After a Negative MP Shock



Moreover, the differential response of old firms markups upon a negative MP shock in the model peaks at a value of 0.6%, while empirically it goes up to 3%. Importantly then, our quantitative framework is able to replicate 20% of the excess counter-cyclicality of old firms' markups to MP shocks that we have estimated in Compustat. In turn, this represent a satisfactory quantitative validation of our framework, which is hence able to replicate both qualitatively and quantitatively the heterogeneity in the response of markups by firm age to MP shocks that has been documented in the data.

5.2 Decomposing the Differential Response of Markups

Analytical Result under Flexible Prices. In what follows, we show that the combination of the total derivatives of the demand function and the desired markup respectively gives the opportunity to understand the heterogeneous response of firm prices to changes in aggregate output Y, wage W and the demand index $\mathcal{D}$. For the sake of analytical tractability, we first carry out such decomposition in a version of the model without price adjustment costs. As derived in Section 3, the demand function in our theoretical framework is given by:

$$y = \left(1 - \omega \log\left(\frac{\sigma}{\sigma - 1}\frac{1}{a}\frac{p}{\mathcal{D}}\right)\right)^{\sigma/\omega} \frac{Y}{a}$$

while the desired markup can be written as follows:

$$\frac{\alpha p}{W y^{\frac{1}{\alpha}-1}} = \frac{\sigma \left(\frac{y}{Y}a\right)^{-\omega/\sigma}}{\sigma \left(\frac{y}{Y}a\right)^{-\omega/\sigma} - 1} \equiv \mu(a)$$

where $\mu(a)$ denotes the markup and increases in the demand faced by the firm. From these two equations, we can derive a set of expressions linking the change in firm prices (and similarly markups) to changes in aggregates W, Y and $\mathcal{D}$ and model parameters (see the derivations in the

Appendix):

$$\frac{\partial \log p}{\partial \log Y} = \frac{\frac{1}{\alpha} - 1}{1 + \left(\frac{1}{\alpha} - 1\right) \frac{\mu(a)}{\mu(a)-1} + \frac{\omega}{\sigma}\mu(a)}$$
$$\frac{\partial \log p}{\partial \log W} = \frac{1}{1 + \left(\frac{1}{\alpha} - 1\right) \frac{\mu(a)}{\mu(a)-1} + \frac{\omega}{\sigma}\mu(a)}$$
$$\frac{\partial \log p}{\partial \log \mathcal{D}} = \frac{\left(\frac{1}{\alpha} - 1\right) \frac{\mu(a)}{\mu(a)-1} + \frac{\omega}{\sigma}\mu(a)}{1 + \left(\frac{1}{\alpha} - 1\right) \frac{\mu(a)}{\mu(a)-1} + \frac{\omega}{\sigma}\mu(a)}$$

where the standard CES equivalent, in the case of perfect competition, can be obtained setting the Kimball superelasticity parameter $\omega = 0$. Notice that all derivatives are positive, which means that the negative MP shock negatively affects W, Y and $\mathcal{D}$. Moreover, the first two derivatives are decreasing in $\mu(a)$ and the third one increases in $\mu(a)$. The same observations hold true if we were to write the derivative of firm's markup with respect to Y, W and $\mathcal{D}$. At the same time, the second derivatives with respect to the demand faced by the firm are given by:

$$\frac{\partial^2 \log p}{\partial \log Y \partial a} < 0, \quad \frac{\partial^2 \log p}{\partial \log W \partial a} < 0, \quad \frac{\partial^2 \log p}{\partial \log \mathcal{D} \partial a} > 0$$

The signs of these second derivatives imply that the prices of firms facing higher demand decline less after the MP shock due to the effects coming from the decline in W and Y, whereas prices decline more due to the decline in $\mathcal{D}$. The aggregate effect prevailing in GE will then depend on the specific parametrization. It is important to stress that, while these derivatives have been taken with respect to the demand a face by firms, there is a strong correlation and direct mapping between the accumulation of demand and firm age progression. This ensures that we can safely interpret the above results as the effects of the changes in Y, W, $\mathcal{D}$ after a MP shock on the prices of relatively older firms.

Benchmark Economy. The same decomposition is then carried out in practice in the quantitative model with nominal rigidities. In particular, we first compute numerically the general equilibrium response of the economy to a negative MP shock. Then, taking as given the equilibrium paths for the aggregate variables $Y, W, \mathcal{D}, r, \pi$ we look at the partial responses of old firms' markups to each of the shocks separately. Before commenting on the quantitative results, we follow the same spirit as in Kaplan et al. (2018) and provide intuition for the channels at play in our fully-fledged economy with heterogeneous firms and endogenous markups. Let us first write the difference between the average markups of old firms and the average markups of young firms as a function of the equilibrium prices, quantities, and inflation. We collect these terms in the vector $\{\mathcal{S}_t\}_{t\geq 0}$, with $\mathcal{S}_t = \{r_t, W_t, Y_t, \mathcal{D}_t, \pi_t\}$, and define the above-mentioned difference $\widehat{\mathcal{M}}(\{\mathcal{S}_t\}_{t\geq 0})$ induced by the path of the monetary shock $\{\varepsilon_t\}_{t\geq 0}$ from its initial hit until it fully reverts to zero as:

$$\widehat{\mathcal{M}}(\{\mathcal{S}_t\}_{t\geq 0}) := \int \mu_t(p, a; \{\mathcal{S}_t\}_{t\geq 0}) \mathbb{1}\{g_t(p, a) \geq \bar{a}\} d\lambda_t - \int \mu_t(p, a; \{\mathcal{S}_t\}_{t\geq 0}) \mathbb{1}\{g_t(p, a) < \bar{a}\} d\lambda_t \tag{27}$$

where $\mu_t(p, a; \{\mathcal{S}_t\}_{t\geq 0})$ is the firm markup, $g_t(p, a)$ is a mapping between firm's states and its age, $\bar{a}$ is the median firms' age, and $d\lambda_t(p, a; \{\mathcal{S}_t\}_{t\geq 0})$ is the joint distribution of prices and idiosyncratic demand. Totally differentiating Equation 27, we decompose the difference in the average markup response between old and young firms at time $t = \tau$ as:

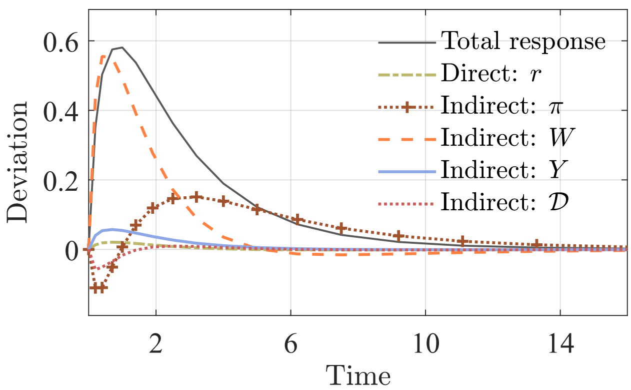
$$d\widehat{\mathcal{M}}_\tau = \underbrace{\int_\tau^\infty \frac{\partial \widehat{\mathcal{M}}_\tau}{\partial r_t} dr_t dt}_{\text{direct effect}} + \underbrace{\int_\tau^\infty \left(\frac{\partial \widehat{\mathcal{M}}_\tau}{\partial W_t} dW_t + \frac{\partial \widehat{\mathcal{M}}_\tau}{\partial Y_t} dY_t + \frac{\partial \widehat{\mathcal{M}}_\tau}{\partial \mathcal{D}_t} d\mathcal{D}_t + \frac{\partial \widehat{\mathcal{M}}_\tau}{\partial \pi_t} d\pi_t \right) dt}_{\text{indirect effect}} \quad (28)$$

where the first term reflects the direct effect of a change in the interest rate, which enters the Euler equation of the agents, holding the other variables of interest constant. The remaining terms in the decomposition reflect the indirect effects of changes in inflation, the real wage, real output and the demand index that arise in general equilibrium after the hit of the MP shock. In practice, we need to compute each of these components numerically. For example, the formal definition of the first term in Equation 28, which is the direct effect of changes in the real interest rate $\{r_t\}_{t\geq 0}$, is:

$$\int_\tau^\infty \frac{\partial \widehat{\mathcal{M}}_\tau}{\partial r_t} dr_t dt = \int_\tau^\infty \frac{\partial \widehat{\mathcal{M}}(\{r_t, \overline{W}, \overline{Y}, \overline{\mathcal{D}}, \overline{\pi}\}_{t\geq 0})}{\partial r_t} dr_t dt. \quad (29)$$

This term is the *partial-equilibrium* response of the difference in the average markups between old and young firm that face a time-varying real interest rate path $\{r_t\}_{t\geq 0}$, but holding the paths for the real wage $\overline{W}$, the real output $\overline{Y}$, the demand index $\overline{\mathcal{D}}$, and nominal inflation rate $\overline{\pi}$ constant at their steady-state values. We calculate this term from the model by feeding these time paths into the firms' (and household's) optimization problem, computing the policy function and their markups for each firms, and aggregating across firms using the corresponding distribution. The other terms in the decomposition are computed in a similar fashion.

Figure 5.2: Decomposing the Differential Response of Markups

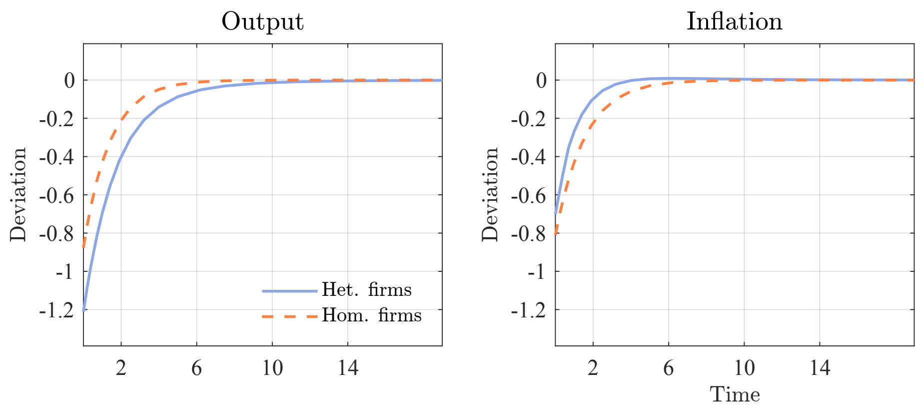


The results of the decomposition exercise are depicted in Figure 5.2. All effects are to be intended as p.p. deviations from the mean response across all firms in the economy. The outer dark line represents the total GE effect of a negative MP shock on the differential impulse response of markups for old firms compared to young ones. Also, note that the GE effect is not a direct sum of the partial effects due to non-linearities in aggregation. Most of the resulting effect on the differential response of old firms' markups is to be attributed to changes in the aggregate W, hence to changes in the cost of labor after a negative shock to the interest rate. Since our model features heterogeneous and endogenous markups in the presence of a Kimball aggregator, big firms (and hence old firms) have a lower passthrough from production costs to prices. In this sense, a negative MP shock in the economy puts a downward pressure on the labor input cost W, but old firms' sales react less than proportionally, as dominant companies do not decrease prices as much. Since markups are the ratio between business sales and costs, the resulting effect on markups is positive, leading to the observed stronger countercyclical response of old firms' markups to a negative MP shock.

5.3 Amplification Mechanism

In what follows, we conclude our quantitative analysis by studying the shock amplification mechanism at work in our economy, comparing our calibrated framework with a standard one-firm NK model. As pointed out in Mongey (2017), in economies where real rigidities are present, shocks have a strong propagation through quantities, which we set to verify in our case. Moreover, we proceed to also explain to which extent both firm heterogeneity and the Kimball aggregator that characterize our model can be responsible for greater swings in macro aggregates after a negative MP shock.

Figure 5.3: Comparing Output and Inflation Responses



To ensure we are working with two comparable economies, we first calibrate the one-firm NK economy to have the same size as in our heterogeneous firms framework (hereafter the FDNK) in terms of overall output produced. Moreover, since the standard NK model features perfect competition, we set the elasticity of substitution σ in its CES aggregator to match an aggregate markup of 1.68, which is the value targeted in the FDNK model under the Kimball aggregator. With the two models at hand, we simulate a negative MP shock and solve for the response of the main macroeconomic aggregates in the two economies. In particular, we analyse the trajectories of inflation π and output Y over an 16-quarters period, and hence compare the relative percentage deviation from steady state values of both prices and quantities. The results of this exercise are depicted in Figure 5.3.

Comparing output and inflation responses across the two models, it is clear that a negative MP shock produces a bigger drop in output and a milder decline in prices in our FDNK set up compared to a standard one-firm NK model. The negative change in the interest rate decreases output by on average 20 p.p. more in the economy characterised by heterogeneous firms and endogenous markups, with the effect lasting for more than 10 quarters after the shock hits. At the same time, prices and hence inflation drop by relatively more in the one-firm NK model, which implies that the presence of the Kimball aggregator and the differential passthrough that characterize our model economy mitigate the downward pressure exerted by the negative MP shock on firm prices.

On the one hand, as argued in Klenow and Willis (2016), the presence of the Kimball aggregator adds a source of real frictions in the NK model, represented by a higher degree of concavity in the firm's profit function with respect to its relative price. Under the Kimball aggregator, sellers face a price elasticity of demand that is increasing in their good's relative price. For instance, when a repricing producer faces lower labor costs after a negative MP shock, it will temper its price drop because of the endogenous increase in its desired markup, and this effect would be stronger the lower the elasticity of demand faced by the producer. Since the presence of a real rigidity makes firms more reluctant to change prices, firms do not pass marginal cost shocks as fully onto their prices as they would in a standard NK model with a CES aggregator. Hence, in our FDNK set up, MP shocks propagate more through quantities than prices, and decrease aggregate output by relatively more.

On the other hand, without heterogeneity on the firm's side, the presence of the Kimball aggregator alone does not automatically imply the amplification of shocks in our setup: in fact, the effects of the real rigidity introduced by the Kimball aggregator kick in only when businesses are indeed heterogeneous and hence characterized by different passthroughs from costs to prices with respect to one another. If all firms were to be equal (as in the representative-firm NK model), they would also be equal to the average firm in the economy and have identical sales shares. Specifically, focusing on Equation 19, the elasticity of demand faced by producers would not vary across firm, and their response to MP shocks would be identical. On the contrary, in our FDNK set up, since big firms (hence old firms) respond more countercyclically than small ones and decrease their prices by less, the propagation of a negative shock gets strengthened. Hence, the heterogeneity of firms, combined with the real rigidity introduced by the Kimball aggregator set up, delivers the amplification mechanism at work in the present model.

6 Conclusion

In this paper, we have taken an empirical and theoretical approach to the study of firm heterogeneity in the response of markups to MP shocks. In order to carry out our data analysis, we have merged exogenously-identified monetary policy shocks series with a rich quarterly dataset comprising publicly-listed companies based in the US between 1990Q1 and 2016Q4. Next, we have documented that old firms' markups tend to increase after a monetary policy tightening, while young firms' markups show a mildly procyclical behavior after a negative interest rate shock. Moreover, our empirical investigation seems to also suggest that the differential response of markups by firm's age could be related to the accumulation of customers and demand over time, which enables older firms to change by relatively less their prices thanks to an established position in their markets.

In our quantitative analysis, we have embedded our findings into a NK model, augmented with heterogeneous firms and a process of demand accumulation, and in which markups arise endogenously and evolve over the life-cycle of the companies. Our calibrated framework can replicate the life-cycle profile of firms' markups and growth rates, and the distribution of companies and employment shares by corporate age. Moreover, we were able to explain up to a fifth of the empirically estimated excess counter-cyclicality in the markups of firms above the median age after a negative monetary policy shock. Finally, we have shown that both firms' heterogeneity and endogenous markups generate amplification in the response of aggregate quantities to contractionary interest rate movements, which further distinguishes our set up from standard frameworks with nominal rigidities. In the future, we aim to further study optimal monetary policies in the presence of imperfect competition, demand accumulation, and heterogeneity in the passthrough from costs to prices.

Appendix

A Data Appendix

Figure A.1: Alternative Specification for Corporate Age

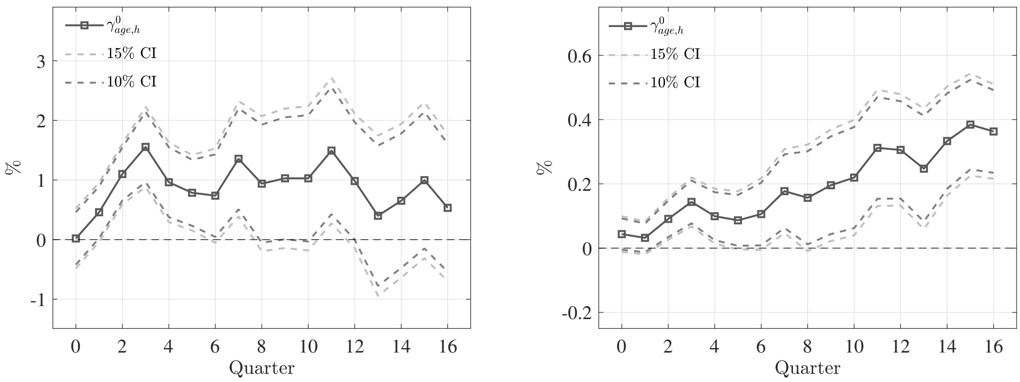


Figure A.2: Excluding Future Shocks (left) and Sector-Quarter FE (right)

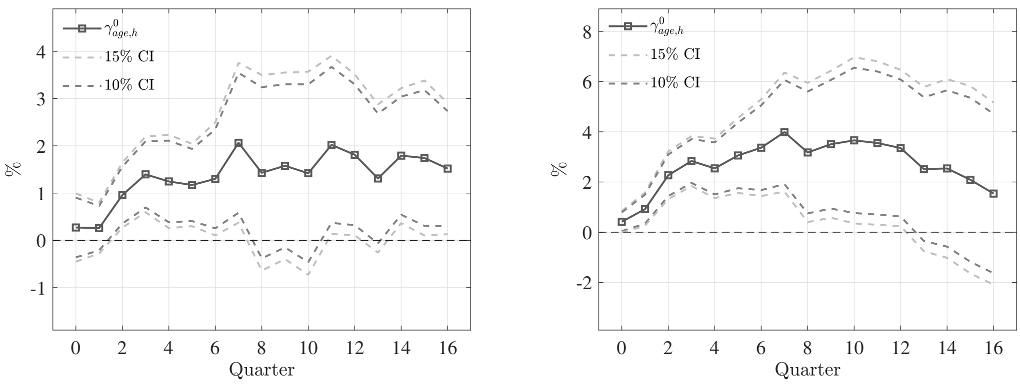


Figure A.3: Age and Leverage

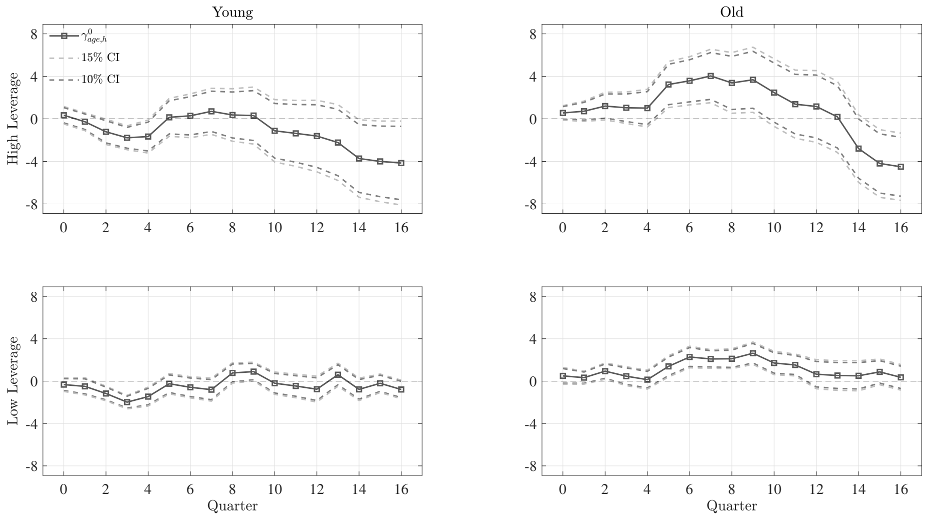


Figure A.4: Age and Liquidity

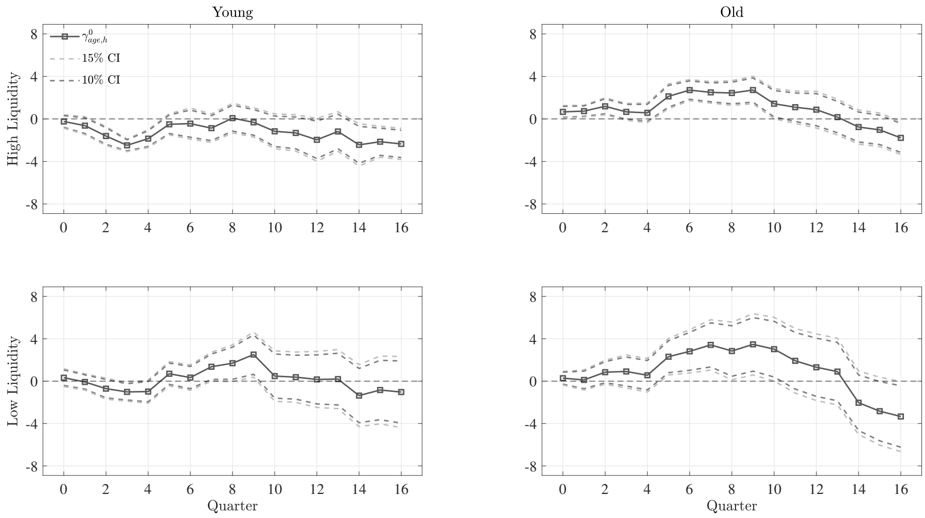
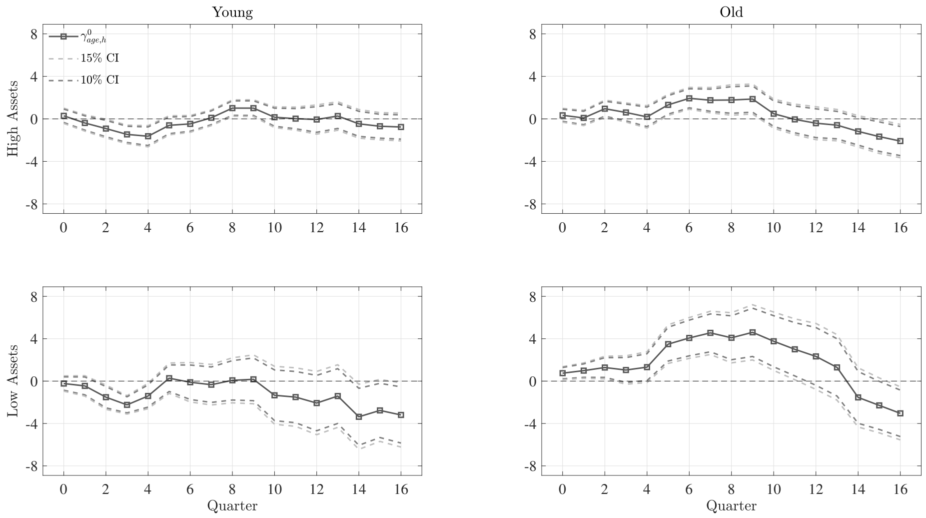


Figure A.5: Age and Size



B Quantitative Appendix

B.1 Decomposition Exercise: Derivations

The demand function in our model is given by:

$$y = \left(1 - \omega \log\left(\frac{\sigma}{\sigma-1}\frac{1}{\xi(a)}\frac{p}{\mathcal{D}}\right)\right)^{\sigma/\omega}\frac{Y}{\xi(a)}$$

Which has the total derivative:

$$dlog\frac{y}{Y}\xi(a) = -\frac{\sigma(dlogp - dlog\mathcal{D})}{1-\omega\log\left(\frac{\sigma}{\sigma-1}\frac{1}{\xi(a)}\frac{p}{\mathcal{D}}\right)} = -\sigma\left(\frac{y}{Y}\xi(a)\right)^{-\omega/\sigma}(dlogp - dlog\mathcal{D})$$

The desired markup is instead defined as follows:

$$\frac{\alpha p}{Wy^{\frac{1}{\alpha}-1}} = \frac{\sigma\left(\frac{y}{Y}\xi(a)\right)^{-\omega/\sigma}}{\sigma\left(\frac{y}{Y}\xi(a)\right)^{-\omega/\sigma}-1}$$

By taking the total derivative it is possible to get:

$$\frac{\alpha}{Wy^{\frac{1}{\alpha}-1}}dp - \frac{\alpha p}{Wy^{\frac{1}{\alpha}-1}}dlogW + \left(1-\frac{1}{\alpha}\right)\frac{\alpha p}{Wy^{\frac{1}{\alpha}}}dy =$$
$$-\frac{1}{\left(\sigma\left(\frac{y}{Y}\xi(a)\right)^{-\omega/\sigma}-1\right)^2}\sigma\left(-\frac{\omega}{\sigma}\right)\left(\frac{y}{Y}\xi(a)\right)^{-\omega/\sigma-1}d\frac{y}{Y}\xi(a)$$

Substituting in the above expression $dy = \frac{Y}{\xi(a)} d\frac{y}{Y}\xi(a) + \frac{y}{Y}dY = y(dlog\frac{y}{Y}\xi(a) + dlogY)$ it is possible to obtained the following equation:

$$\frac{\alpha p}{Wy^{\frac{1}{\alpha}-1}} dlogp - \frac{\alpha p}{Wy^{\frac{1}{\alpha}-1}} dlogW + \left(1 - \frac{1}{\alpha}\right) \frac{\alpha p}{Wy^{\frac{1}{\alpha}-1}} (dlog\frac{y}{Y}\xi(a) + dlogY) =$$
$$-\frac{1}{\left(\sigma\left(\frac{y}{Y}\xi(a)\right)^{-\omega/\sigma} - 1\right)^2} \sigma\left(-\frac{\omega}{\sigma}\right)\left(\frac{y}{Y}\xi(a)\right)^{-\omega/\sigma} dlog\frac{y}{Y}\xi(a)$$

which in turn implies:

$$dlogp - dlogW + \left(1 - \frac{1}{\alpha}\right)(dlog\frac{y}{Y}\xi(a) + dlogY) = \frac{1}{\left(\sigma\left(\frac{y}{Y}\xi(a)\right)^{-\omega/\sigma} - 1\right)}\left(\frac{\omega}{\sigma}\right) dlog\frac{y}{Y}\xi(a)$$

$$dlogp - \left(\left(\frac{1}{\alpha} - 1\right) + \frac{1}{\left(\sigma\left(\frac{y}{Y}\xi(a)\right)^{-\omega/\sigma} - 1\right)}\left(\frac{\omega}{\sigma}\right)\right) dlog\frac{y}{Y}\xi(a) = dlogW + \left(\frac{1}{\alpha} - 1\right) dlogY$$

Substituting $dlog\frac{y}{Y}\xi(a)$ from the total derivative of the demand function we get:

$$dlogp + \left(\left(\frac{1}{\alpha} - 1\right) + \frac{1}{\left(\sigma\left(\frac{y}{Y}\xi(a)\right)^{-\omega/\sigma} - 1\right)}\left(\frac{\omega}{\sigma}\right)\right)\sigma\left(\frac{y}{Y}\xi(a)\right)^{-\omega/\sigma}(dlogp - dlog\mathcal{D}) =$$
$$dlogW + \left(\frac{1}{\alpha} - 1\right) dlogY$$

$$dlogp + \left(\left(\frac{1}{\alpha} - 1\right)\frac{\mu(a)}{\mu(a) - 1} + \frac{\omega}{\sigma}\mu(a)\right)(dlogp - dlog\mathcal{D}) = dlogW + \left(\frac{1}{\alpha} - 1\right) dlogY$$

$$dlogp = \frac{dlogW + \left(\frac{1}{\alpha} - 1\right) dlogY + \left(\left(\frac{1}{\alpha} - 1\right)\frac{\mu(a)}{\mu(a)-1} + \frac{\omega}{\sigma}\mu(a)\right) dlog\mathcal{D}}{1 + \left(\frac{1}{\alpha} - 1\right)\frac{\mu(a)}{\mu(a)-1} + \frac{\omega}{\sigma}\mu(a)}$$

Where $\mu(a)$ is the markup. The above expression can be rearranged to get the relative contributions of Y, W and D to the change in prices and markups, reported in the main text.

Part III

Labor and Family Dynamics in a Joint-Search Framework

Marta Morazzoni* and Danila Smirnov†

1 Introduction

In the US, 60% of labor force participants are married,[1] and available empirical evidence suggests the presence of stark heterogeneities in wages and unemployment rates by marital status. How do marital dynamics influence the labor market performance of individuals? Given their disparities in labor market outcomes, should optimal policies target differently married and single households? In this project, we build on the seminal contribution by Guler et al. (2012) and propose a novel quantitative framework to tackle this set of questions.[2] Specifically, we endogenize family formation in a joint-search model and show that agents' selection and sorting in the marriage market are key to quantitatively replicate the observed differences in wages and unemployment rates by marital status. Moreover, we establish that the interaction of marital and labor market decisions has important implications for the design of optimal unemployment benefits.

Using US data from the Current Population Survey (CPS) and the 1979 National Longitudinal Survey of Youth (NLSY 79), we first document that married agents have on average 20% higher wages and 2.5 percentage points lower unemployment rates compared to singles of similar characteristics. These results highlight the existence of a *wage marital premium* (WMP) and an *unemployment marital gap* (UMG), consistent with the evidence in Choi and Valladares-Esteban (2018). Second, we show that part of the estimated labor market differences between single and married households is unconditional on being married. In fact, the married sample has a larger fraction of highly educated agents, who tend to have better labor market outcomes and to partner with individuals of similar educational attainments. This suggests that agents sort and select into marriage, which in turn can contribute to differences in wages and unemployment rates by marital status.

To rationalize these findings, we build and quantify a novel heterogeneous agents model in continuous time that combines together endogenous marriage and job search. In our framework, individuals differ in their labor productivity, which determines wages in efficiency units. Productivity

*Department of Economics and Business, Universitat Pompeu Fabra. Email: marta.morazzoni@upf.edu. We are indebted to Isaac Baley, Davide Debortoli and Edouard Schaal for their constant help and guidance. We thank Matthias Doepke, Nezih Guner, Fatih Guvenen, Stefania Marcassa, Bastian Schulz, Andrea Sy, Arnau Valladares-Esteban and Jaume Ventura for helpful comments, as well as participants in the CREI Macro Lunch, the CERGY Webinar in Gender and Family Economics, the 4th Dale T. Mortensen Conference, the 2021 European Economic Association, the 2021 Spanish Macroeconomic Network, the 2022 Royal Economic Society and several other workshops for invaluable discussions. We thankfully acknowledge financial support from the Spanish Ministry of Economy and Competitiveness through the Severo Ochoa Programme for Centres of Excellence in R&D (CEX2019-000915-S).

†Department of Economics and Business, Universitat Pompeu Fabra. Email: danila.smirnov@upf.edu

[1] Male and female agents considered. Note that the labor force participation rate of married women is currently 70%.

[2] See Flabbi and Mabli (2018) on the quantitative fit of joint-search models with respect to labor market differences by marital status, Flabbi and Finn (2015) on the interplay between education, household's and labor market decisions, and Gemici (2011) on the connection between migration, within-household bargaining and labor market outcomes.

is accumulated on the job and depleted off the job, and it is subject to uninsurable idiosyncratic shocks. With respect to labor market outcomes, both single and married agents within couples can be either employed, unemployed, or out of the labor force. As in McCall (1970), job search is random and unemployed households receive wage offers at an exogenous rate, while enjoying publicly-funded unemployment benefits. When working, they consume their wage in efficiency units and net of labor income taxes, and face the risk of losing their job. Agents can also decide to leave the labor force and enjoy home production. Importantly, the working life of all households in our model spans on average 40 years before they retire and exit the economy.

Regarding the dynamics of the marriage market, the main theoretical contribution of the paper is to allow the choice of marrying to be endogenous and bilateral. Specifically, all individuals in our model economy are born single and face an exogenous probability of meeting partners throughout their life time, who are characterized by heterogeneous labor productivities and employment outcomes. Upon meeting potential partners on the marriage market, agents assess the returns of matching compared to staying single and continuing searching for a spouse. When married, households split the sum of their respective incomes – either from working or from being unemployed – and behave as a joint-searcher on the labor market. In addition, marriage increases the probability of realized fertility and hence the likelihood of bearing child-related expenses.

Two main mechanisms are at work in the model. On the one hand, agents have strong incentives to marry in order to share their respective incomes and benefit from the informal insurance provided within households against labor productivity shocks and unemployment risks. Pooling income shields individuals from large drops in consumption upon negative labor market outcomes, and allows them to wait longer for a better wage offer when searching for a job. On the other hand, being in a couple can be costly if a spouse suffers poor labor market outcomes and because joint-households more often bear child-related expenses. These elements foster ex-ante selection into marriage based on labor productivity and labor market outcomes. Furthermore, the assumption that both parts have to agree on marrying consistently generates positive sorting among partners. Ex-post, selection and sorting imply that married households have on average higher labor productivity, higher wages and lower unemployment rates compared to singles.

We then calibrate the model on US data: a novel solution method we develop allows to conveniently compute 12 endogenous households' distributions over labor productivity, marital status and labor market outcomes. The calibrated framework replicates – as untargeted moments – 75% and 50% of the differences in wages and unemployment rates across married and single households in the US. It is important to stress that explaining both the WMP and the UMG is the main contribution of our quantitative exercise. In particular, we estimate that a model abstracting from sorting and selection into marriage can explain only half of the empirically observed WMP. This is due to the fact that the informal insurance provided within couples affects the reservation wage strategy of joint-searchers. Searching longer for better job offers is consistent with married individuals having higher salaries, but it is counterfactual with respect to their lower unemployment rates. Accounting for labor productivity differences across the samples of singles and married, which are endogenously generated by agents' sorting and selection into marriage, can instead reconcile theoretical predictions and empirical observations regarding both the WMP and the UMG.

Next, we assess the performance of our quantitative framework with respect to several marital patterns documented for the US, which are also not targeted in the calibration. In the model, households tend to marry when they are 31 years old, consistent with an average age at first marriage of 29 years old for the US over the past decade. Second, we replicate nearly 70% of the positive assortative matching of agents within couples based on their labor market outcomes. As a final remark, the model reproduces 80% of the empirically estimated semi-elasticity of marriage rates to individuals' educational attainment, which we proxy in our model by agents' initial labor productivity. Hence, allowing for endogenous marriage within a theory of joint-search in the labor market contributes to explaining not only the differences in wages and unemployment rates by marital status, but also several patterns of sorting and selection into couples observed in US data.

Finally, since our framework features a government that collects labor income taxes to fund unemployment insurance schemes, we use the model as a quantitative laboratory to study the optimal design of unemployment benefits. Note that all workers in the economy are subject to labor productivity shocks and unemployment risk, and adverse labor market outcomes cause sizable falls in households' income and consumption. However, due to resource-sharing within couples, married individuals can count on the income of their spouses to mitigate consumption losses. Being in a joint-household also allows agents to wait relatively longer before accepting a better job offer. Consistent with these observations, we find that it would be optimal to redistribute a share of fiscal resources from married to single households. This is achieved by increasing the unemployment benefit of single agents, while keeping fixed our baseline labor income tax rate.

On the other hand, all individuals in the economy face the probability of realized fertility and subsequent child-related expenses, which is higher for agents in couples compared to singles. This feature of the model introduces a second tradeoff in the design of optimal transfers, which could in principle lead to a redistribution of fiscal resources towards married agents as opposed to singles. In fact, when considering both the marital and parental status of recipients, we show that an optimal unemployment insurance scheme should redistribute resources to single parents. Such optimal policy can achieve a 1% welfare gain with respect to the baseline calibration. Our analysis is hence a first attempt to disentangle how to balance the provision of both public unemployment insurance and child-related benefits, especially considering the family status of recipients.

Related Literature. Our paper builds on the joint-search framework developed by Guler et al. (2012), but allows individuals to differ by labor productivity, marital and fertility outcomes to study the interplay of family formation and job search decisions. Our work is also complementary to Pilossoph and Wee (2021), who analyze the WMP in the US as the result of higher reservation wage strategies for married households. In our paper, we however model endogenous marriage to explain both the WMP and the UMG,[3] and show that the the sorting and selection of individuals into joint-households is key to replicate the difference in wages and unemployment rates by marital status observed in the US. Moreover, we are similar in spirit to Calvo et al. (2021), who build a model of sorting in the marriage and labor markets to investigate how complementarities in home-production hours across spouses affect income inequality within and between households.

Our work is also related to several studies that have analyzed the interaction of households' dynamics and labor market outcomes focusing on the *intensive* margin of labor supply, as in Ortigueira and Siassi (2013), Guner et al. (2012), Greenwood et al. (2014), Doepke and Tertilt (2016) and Pilossoph and Wee (2020). We instead focus on the *extensive* margin of labor supply and explore the interplay of marriage and job search choices. In so doing, we provide an assessment of which workers tend to marry, whom they form a family with, how early in their career they typically marry and how this choice affects job search strategies, wages and unemployment rates.

Moreover, our paper contributes to a growing literature that analyzes fiscal policies in models of heterogeneous households with different marital status and family characteristics. For example, Guner et al. (2012), Borella et al. (2019) and Ortigueira and Siassi (2020) investigate the impact

[3]See the evidence reported first in Choi and Valladares-Esteban (2018) and McConnell and Valladares-Esteban (2020).

of joint taxation schemes on the labor market outcomes of singles and married workers, focusing on the intensive margin of labor supply, while Guner et al. (2020) explores the macroeconomic effects of child benefits. In a different spirit, Choi and Valladares-Esteban (2019) study optimal unemployment insurance in an incomplete markets model with single and married households, which however abstracts from endogenous marital dynamics.[4] The key contribution of our paper is to endogenize both labor market and family decisions, and analyze how sorting and selection into couples affect the optimal provision of unemployment and child-related benefits. In particular, we examine how government transfers can benefit agents that have no access to the informal insurance provided by the family against labor market and productivity shocks, and how this margin of redistribution is further affected by the presence of children within the household.

2 Empirical Evidence

The following section discusses several labor and marriage markets facts for the US civilian non-institutional population, both in the cross-section and over time. We document the presence and extent of a wage marital premium (WMP) and an unemployment marital gap (UMG) in the data. Then, we provide suggestive evidence on the selection of highly-educated individuals into joint-households, which could partially explain the observed labor market differences by marital status.

2.1 Wage and Unemployment Differences by Marital Status

To empirically analyze the wage and unemployment rate of married and single individuals, we use CPS, a repeated annual cross-section of 65,000 US households, and focus on the 2000-2019 period. The sample contains details on individual characteristics such as gender, ethnicity, age, marital status, number of children, and education, among other demographic factors. When applicable, it also includes a section on the respondent's spouse, as well as information on labor force participation, employment status, salaried income, hours worked, and 4 digits occupational code.

We restrict our focus to single and married individuals whose age is between 18 and 65 years old. In any given sample year, we define as "singles" those agents who never married before, while we consider as "married" respondents that are legally married.[5] In terms of labor market outcomes, we count as "out of the labor force" all the individuals that are in the working age population but are neither employed nor unemployed, excluding students, retirees and agents with disabilities. Moreover, we define as "yearly" unemployed those individuals reporting 52 weeks in unemployment over a year, and whenever focusing on employed agents we exclude those with no data on hourly earnings. We then use our final CPS sample to run the following OLS regression specifications:

$$y_{it} = \beta_0 + \beta_1 \mathbb{1}_{married} + \delta' \mathbf{\Gamma}_{it} + \alpha_t + \eta_{o(it)} \mathbb{1}_{employed} + \varepsilon_{it} \tag{30}$$

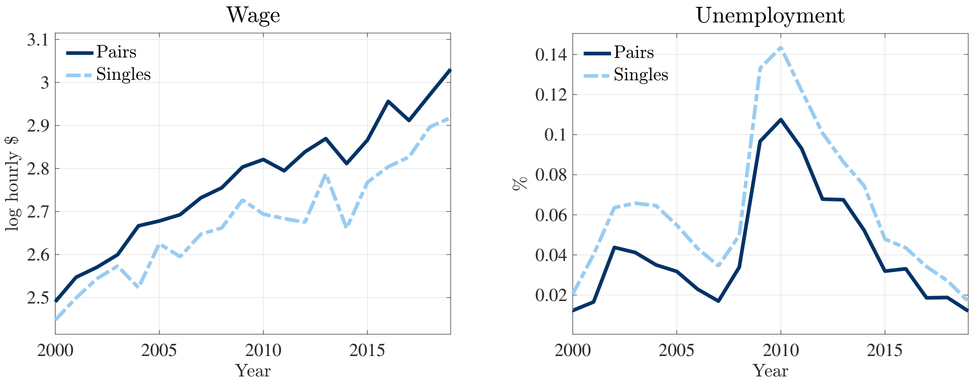
where $y_{it} = \{wage_{it}, unemp_{it}\}$. The key explanatory variable is $\mathbb{1}_{married}$, a dummy variable that is equal 1 if individuals are married and to 0 if they are single. The regressions include a set of controls $\mathbf{\Gamma}$ that captures factors that may affect the income of agents, such as their gender, education, ethnicity, age, number of children and labor market status in the previous year. We

[4] Ek and Holmlund (2010) also study optimal unemployment insurance of agents in couples using a Diamond-Mortensen-Pissarides framework without individual heterogeneity and endogenous household's formation.

[5] Results are robust to define as "married individuals" also co-habitating households, and to count divorced, separated and widowed agents as singles. We can further restrict the sample to include only 25-65 years old individuals.

also allow for year (α_t) and occupation ($\eta_{o(it)}$) fixed effects. For married agents, we include the education and current labor market status of the spouse. Survey weights are used to ensure representativeness.

Figure 2.1: Wage and Unemployment Marital Differences



Data: US Current Population Survey, 2000-2019. Survey weights are used. Controls: gender, age, ethnicity, number of kids, education, employment status in previous year, year and occupation fixed effects, spousal education and employment status.

As reported in Table A1, being married is on average associated with having 10% higher hourly wages, 20% higher yearly wages, and 2.5 percentage points lower unemployment rates. For a visual reference, Figure 2.1 plots the average (log) hourly wage and unemployment rate of married and singles as estimated in Equation 32 under our set of control variables. Throughout the years considered, married is associated with having both higher wages and lower unemployment rates. Moreover, marital differences in both (log) hourly wages and unemployment rates have remained particularly stable over time. Our results are qualitatively and quantitatively in line with the findings of Choi and Valladares-Esteban (2018).[6]

To explain the WMP it has been argued that, since being in a couple can provide informal insurance against labor market risks, married unemployed agents can wait longer before accepting a job offer, which could ex-post rationalize their higher wages. Using a joint-search model à la Guler et al. (2012), Pilossoph and Wee (2021) have shown that married agents' higher reservation wage strategy can determine up to 16-41% of the estimated WMP for men, and between 20-68% for women. However, such mechanism alone has a counterfactual prediction on the unemployment rates of single and joint-households. If married agents can wait longer to accept a job, they should have higher unemployment rates, contrary to what observed empirically. Hence, at least part of the observed differences in labor market outcomes seems to be *unconditional* on marrying. In what follows, we explore *selection* into marriage on ex-ante heterogeneous ability, that we proxy by educational attainment, as a plausible factor influencing the labor market performances of married and single individuals, and argue that such channel can be consistent with both the WMP and UMG observed in US data.

[6] Hill (1979) is the first study to show that married men have higher average wages than single men. Other authors have further documented a consistent WMP across different US samples of male individuals (see for instance Korenman and Neumark (1991) and Antonovics and Town (2004)), whereas recent evidence has stressed the existence of a WMP also between married and single women (see Juhn and McCue (2016), and McConnell and Valladares-Esteban (2020)).

2.2 Selection and Sorting into Join-Households

If the pool of married agents is inherently different from the pool of single households in some key dimension that also affect individuals' labor market outcomes, *selection* into joint-households can play an important role in explaining both the WMP and the UMG. To highlight the ex-ante differences across married and single individuals, we first show that both current and future individuals' marital status correlate with higher wages and lower unemployment rates. To this end, we turn to the 1979 National Longitudinal Survey of Youth (NLSY79), a panel dataset of nearly 13,000 individuals followed since 1979 until nowadays. NLSY79 covers less individuals than CPS, but, in this specific case, it allows us to exploit its panel dimension and consider the evolution of agents' marital status over time. We again restrict our attention to individuals between 25 and 65 years old and on their (log) hourly wage ($wage_{it}$) and employment status ($unemp_{it}$) in each period. Then, we run the following two sets of OLS regression specifications:

$$\{wage_{it}, unemp_{it}\} = \beta_0 + \beta_1 \mathbb{1}_{married} + \delta' \mathbf{\Gamma}_{it} + \alpha_t + \varepsilon_{it} \tag{31}$$

$$\{wage_{it}, unemp_{it}\} = \beta_0 + \beta_1 \mathbb{1}_{married\ next\ year} + \delta' \mathbf{\Gamma}_{it} + \alpha_t + +\varepsilon_{it} \tag{32}$$

In light of the previous discussion, we focus on two key explanatory variables that can affect individual wages and the likelihood of being unemployed: first, $\mathbb{1}_{married}$ is an indicator that takes on a value of 1 if individuals are married and 0 if they are single. And secondly, $\mathbb{1}_{married\ next\ year}$ is an indicator equals to 1 if individuals will marry within one year from current period t, and to 0 if they will remain singles. All the regressions include a set of controls $\mathbf{\Gamma}$ that captures factors that may affect the income of agents, such as their gender, education, ethnicity, number of kids, age, labor market status in the previous year and an indicator for whether the respondent lives in a urban or rural area. We also allow for year fixed effects (α_t).

Table 1: Labor Market Outcomes in NLSY79

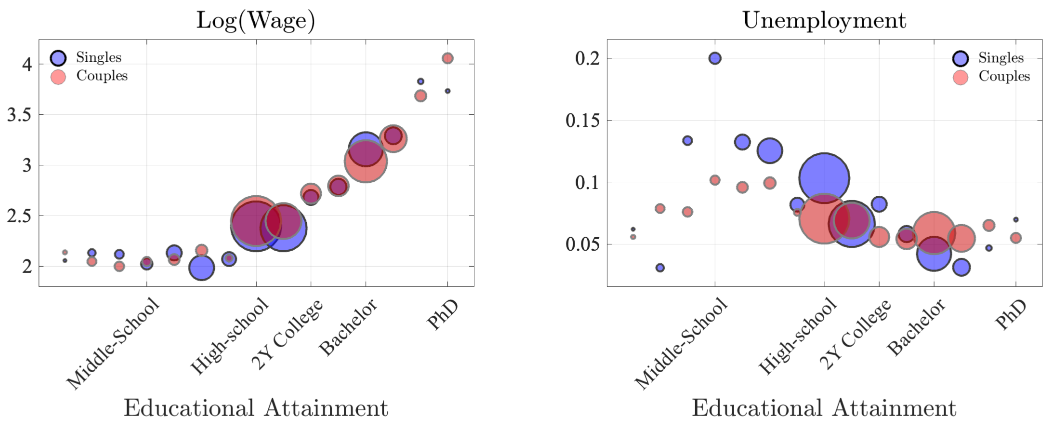
	hourly wage	unemployment	hourly wage	unemployment
Married	0.1240***	-0.0203***		
	(0.0066)	(0.0024)		
Married Next Year			0.1117***	-0.0146**
			(0.0168)	(0.0062)
Controls	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Observations	66,420	83,206	18,266	20,422
R^2	0.3621	0.0816	0.2510	0.1220

Notes: Robust stadard errors in parentheses. Survey weights are used. ***p<0.01, **p<0.05, *p<0.1. Control variables include gender,education, weeks worked in the previous calendar year, an urban vs rural area indicator, ethnicity, kids, and age.

As reported in columns (1) and (3) of Table 1, the estimated WMP and UMG in NLSY79 are consistent with the ones documented in CPS, and indicate that being married is correlated with having higher wages and a lower likelihood of being unemployed. However, columns (2) and (4) illustrate that singles agents who marry in $t+1$ already show higher wages at time t compared to those who remain singles, as well as a lower probability of being unemployed. This suggests that agents who eventually marry tend to have better labor market outcomes to begin with and unconditional on marrying. One possible explanation is that the empirically observed differences in wages and unemployment rates across married and singles individuals are partially due to the way agents select and sort into marriage.[7] If the most productive individuals were the ones that effectively marry, married households would display lower unemployment rates and higher wages, as unemployment decreases and wages increase in agents' productivity. In this case, the WMP and the UMG could be partially a result of the fact that married are more represented among the most productive individuals, which are also the agents with better labor market outcomes. Under the assumption that educational attainment does not fully predict households' income, the unobserved variation in productivity can not be filtered by the regression controls, and thus could contribute to the difference in wages and unemployment between married and single individuals.

To test further our hypothesis, we leverage available information on individuals' education in CPS to compute (log) wages and unemployment rates by educational attainment for married and single agents.[8] We then regress both labor market outcomes using the same controls as in Figure 2.1 and plot the estimation residuals in Figure 2.2. Note that each circle is proportional to the share of married and singles within a given educational category. Two observations can be made: on the one hand, wages on average increase in educational attainment of the households (our proxy for agent's productivity). Similarly, unemployment rates decline for highly educated individuals, suggesting that the higher the education of a person, the better her labor market outcomes.

Figure 2.2: Selection Effect on Wages and Unemployment in CPS (1998-2018)



On the other hand, the size of the circles in the graphs points out that the share of married among highly educated individuals is higher. Since married agents are more represented among highly educated individuals, whose labor market outcomes are generally better, the different productivity composition of the single and married sample of households seems an important driver of their wage and unemployment differences. This mechanism would be further enhanced if agents were to sort in couples along individual productivities (in the data: educational attainments) and hence be married to partners of potentially similar labor market outcomes (see for example Greenwood et al. (2014)). According to this line of reasoning, marital sorting and selection seem crucial *endogenous* channels to model in order to reconcile the empirical evidence on both the WMP and the UMG with existing theoretical frameworks of joint-search.

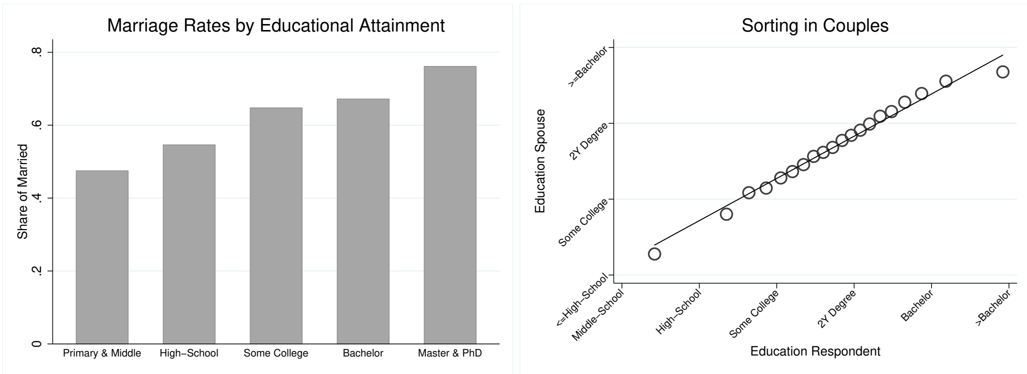
As a final consideration, we provide additional suggestive evidence of marital selection and sorting in the sample of CPS couples. We first classify agents by their educational attainment, and compute the share of married for each educational grade. To group consistently observations, we create the following educational categories: elementary and middle school, up to high-school diploma, some college (including 2 years associate degrees), 4-6 years of college (including Bach-

[7]See Nakosteen and Zimmer (1987) for a first contribution to this debate.

[8]A similar result holds in NLSY79. We use here CPS due to its larger sample size and higher representativeness.

elor), master and PhD. The left panel of Figure 2.3 shows that marriage rates are increasing in the educational attainment of individuals, as documented in Doepke and Tertilt (2016). We also estimate that the semi-elasticity of marriage rates to an increase in educational attainment is 19%. For instance, focusing on the period between 1998 and 2018, the share of married individuals among master and PhD graduates is 50% bigger than among people with less than a high-school diploma. Moreover, albeit the overall decline in US marriage rates documented in recent decades, 80% of the CPS households in the the right tail of the education distribution marry in their life.

Figure 2.3: Marriage Rates and Sorting in CPS (1998-2018)



Data: US Current Population Survey. The bin scatter plot groups all households in the sample according to their schooling years.

Secondly, the right panel in Figure 2.3 reports the correlation between partners' educational attainments: focusing on the period between 1998 and 2018 and on the households for which we have information on their spouse's education, we document a significant degree of sorting of partners within couples, consistent with previous studies. Considering a more disaggregated measure of education that counts the number of years spent in school instead of grouping observations into broad categories, the correlation between survey respondents' education and the one of their spouses is above 50%, which is a sign of marital sorting. Note that we have focused on the similarity in educational attainments, but an analogous point could be made measuring instead the correlation in wages or personal wealth across spouses.

3 The Model

Our goal is to nest search in the labor market and search in the marriage market together in a tractable heterogeneous agents model. Heterogeneity stems from individual productivities, which are subject to idiosyncratic shocks. Moreover, endogenous choices in both labor and marriage markets further distinguish agents in the economy by their employment and marital outcomes at any time. With respect to job search, we build on McCall (1970) seminal model and also allow agents of any marital status to be either *in* or *out* of the labor force. Households out of the labor force enjoy utility from home production, whereas agents in the labor force can be either employed or unemployed. Employed individuals work for a wage and face exogenous separation shocks, whereas unemployed agents receive unemployment benefits and search for a job.

Moreover, the model features single and joint-searchers as in Guler et al. (2012), and, since we allow for the decision to marry, gives rise to an *endogenous* distribution of married and singles. Agents on the marriage market face a distribution of possible partners – with different productivities and labor market characteristics – and know that marrying entails sharing incomes. The insurance channel provided by the spousal income is an incentive to match, but being in a couple can be costly whenever the partner is hit by negative idiosyncratic and labor market shocks. Moreover, the model allows for exogenous fertility, with a higher probability of arrival for married compared to singles, which increases the likelihood to bear children-related expenses after marrying. Since matching requires bilateral agreement, both channels contribute to sorting on the marriage market, and to selection into couples along individuals' productivities.

3.1 Model Primitives

The model is built over an infinite horizon, continuous time setting. Agents are born single, unemployed and with an initial productivity a_0, which is heterogeneous across households and drawn from a truncated log-normal distribution $\log \mathcal{N}(\mu_a, \sigma_a)$ over the support $[a_{min}, a_{max}]$. Over their life-time, agents endogenously decide to be single or married, as discussed further below. Moreover, all individuals retire according to a Poisson process of rate ξ, and agents in couples retire together.

Idiosyncratic productivity: During their life-time, the individual productivity a of employed and unemployed agents follows the exogenous stochastic processes given by:

$$da_{i,t} = \phi_e(\bar{a} - a_{i,t})dt + \epsilon dW_{i,t} \quad \text{and} \quad da_{i,t} = \phi_u(a_{min} - a_{i,t})dt + \epsilon dW_{i,t} \tag{33}$$

In particular, ϕ_e is the persistence in the productivity process of employed agents, $\bar{a}$ is a common long-term mean productivity, and $\epsilon dW_{i,t}$ is the idiosyncratic innovation term, with dW being a standard Wiener Process. Individuals are able to accumulate skills when employed but, when unemployed, their productivity drifts towards the minimum level of a_{min}, and is characterized by persistence ϕ_u and idiosyncratic shocks $\epsilon dW_{i,t}$. We hence allow individuals to deplete and accumulate $a_{i,t}$ at different rates, which will be pinned down numerically in the quantitative exercise.[9] Finally, we note that our model features (uninsurable) idiosyncratic shocks to individual's productivity, while there is no source of aggregate uncertainty in the economy as a whole.

Preferences: Agents have a concave utility function over consumption, which satisfies standard Inada conditions, and where the coefficient of risk aversion is denoted by γ and assumed to be the same irrespective of the household's marital status.[10] Individuals consume a single consumption good and they discount the future at rate ρ.

Single individuals: When out of the labor force (hereafter outLF), single individuals consume a home-production good h, which we will allow later on in the quantification to differ according to the presence of children in the household.[11] Single individuals in the labor force consume their wage $w(a_{i,t})$ net of labor income taxes τ if they are working, and unemployment benefit $b(a_{i,t})$ if they are not working (both depend on their productivity $a_{i,t}$). Their instantaneous utility is given

[9] In principle, we do not allow for differences in the parameters of the productivity processes by marital status.

[10] Concavity in the utility function is an important feature of joint-search models as it implies differences in the search strategies of single and joint-households. In our framework, single and joint-households' problem are different even under the assumption of linear utilities, but the concavity of the utility function strengthens the incentive to marry.

[11] Home production h differs based on the presence of children in the households. In particular, in the quantitative exercise we allow for $h_k > h_{nk}$ to match empirical differences in the shares of out of the labor force individuals across married and singles, given the differences in their respective likelihoods of having children. Ultimately, allowing for $h_k > h_{nk}$ can be thought as considering childcare provision within the household as part of home production.

by:

$$\text{employed:} \quad u((1-\tau)w(a_{i,t})); \qquad \text{unemployed:} \quad u(b(a_{i,t})); \qquad \text{outLF:} \quad u(h)$$

Married individuals: When out of the labor force, married agents dedicate to home-production h, similarly to singles. When in the labor force, they earn the wage $w(a_{i,t})$ net of labor income taxes τ if they are working, and receive unemployment benefit $b(a_{i,t})$ if they are not working. Regardless of their labor market status, married individuals join their respective incomes, and each agent enjoys half of the surplus (we hence use a unitary model of marriage).[12] Taking a couple formed by individuals i and j, the instantaneous utility for any agent in the pair can be one of the following:

- $u\left(\frac{(1-\tau)(w(a_{i,t})+w(a_{j,t}))}{2}\right)$ if both agent i and j in the couple are working
- $u\left(\frac{(1-\tau)w(a_{i,t})+b(a_{j,t})}{2}\right)$ or $u\left(\frac{(1-\tau)w(a_{j,t})+b(a_{i,t})}{2}\right)$ if only one agent is working
- $u\left(\frac{b(a_{i,t})+b(a_{j,t})}{2}\right)$ if both agent i and j are unemployed
- $u\left(\frac{(1-\tau)w(a_{i,t})+h}{2}\right)$ or $u\left(\frac{(1-\tau)w(a_{j,t})+h}{2}\right)$ if one agent is working and one is outLF
- $u\left(\frac{b(a_{i,t})+h}{2}\right)$ or $u\left(\frac{b(a_{j,t})+h}{2}\right)$ if one agent is unemployed and one is outLF
- $u\left(\frac{h+h}{2}\right)$ if both agent i and j are out of the labor force

3.2 Labor Markets

When in the labor force, agents of any marital status can be employed or unemployed. Job search is random: when unemployed, individuals search and find job offers at rate λ_u. Jobs are characterized by a wage offer, which is drawn from a distribution $f(w)$: the effective wage is given by a combination of the wage draw $\tilde{w}_{i,t}$ and individual's productivity $a_{i,t}$, such that, for an individual of given productivity $a_{i,t}$, the efficiency-unit wage can be written as: $w_{i,t} = \tilde{w}_{i,t}e^{a_{i,t}}$. Intuitively, workers with high productivities enjoy higher wages compared to agents of lower productivities.

Regarding employed workers, we allow matches to end upon exogenous separation, at rate δ. Once a job-worker match breaks up, the worker flows into unemployment and the job disappears.[13]

[12] We assume that agents in a couple equally benefit from their joint income, as it is beyond our scope to analyze joint-surplus splitting rules and bargaining within the household. In fact, if we were to introduce asymmetries in the surplus-sharing of couples, the main interplay between marriage choices and labor market outcomes will remain.

[13] In a different set up, we could allow for employment-to-employment transitions, where matches between a job and a worker can end if agents find a better job opportunity while working. On-the-job search would incentivize both married and singles to quit unemployment faster and climb the job ladder while working. Such mechanism would have a relatively stronger effect on married individuals, who have ex-ante higher reservation wage strategies, further amplifying the unemployment rate gap between married and singles, and strengthening the narrative we have built.

Importantly, agents can leave the labor force at any point during their life, both when employed or unemployed. For the sake of tractability, we assume that being out of the labor force is an absorbing state, and individuals that have exited the labor force cannot rejoin it. Finally, we do not assume job finding and job separation rates to vary across agents of different marital status; we will nonetheless show that the main mechanisms at play in the model – namely within household's insurance and selection into couples – are enough to produce shares of couples and singles across labor market statuses consistent with their data counterparts.

Given the existence of single and joint-searchers in the labor market, there are nine different value functions according to the job and marital status of the households. The value for a married couple composed by agents of given productivities $a_{i,t}$ and $a_{j,t}$ can be one of the following cases:

1. $WW(a_{i,t}, \tilde{w}_{i,t}, a_{j,t}, \tilde{w}_{j,t})$: both agents are employed
2. $WU(a_{i,t}, \tilde{w}_{i,t}, a_{j,t})$ or $UW(a_{i,t}, a_{j,t}, \tilde{w}_{j,t})$: one agent is employed, the other is unemployed
3. $UU(a_{i,t}, a_{j,t})$: both agents are unemployed
4. $WO(a_{i,t}, \tilde{w}_{i,t}, a_{j,t})$ or $OW(a_{i,t}, a_{j,t}, \tilde{w}_{j,t})$: one agent is employed, one is outLF
5. $UO(a_{i,t}, a_{j,t})$ or $OU(a_{i,t}, a_{j,t})$: one agent is employed, the other one is unemployed
6. $OO(a_{i,t}, a_{j,t})$: both agents are outLF

Parallel to that, the value for a single agent of given productivity a_i can be one of the following:

1. $W(a_{i,t}, \tilde{w}_{i,t})$: the agent is employed
2. $U(a_{i,t})$: the agent is unemployed
3. $O(a_{i,t})$: the agent is outLF

Unemployment Benefits: The fiscal surplus collected from the labor income taxes across the employed population is used to finance public insurance for the unemployed population. This feature is equivalent to assuming that there is a public sector whose budget constraint clears at each point in time, so that the sum of collected labor income taxes T and social expenditure G to finance unemployment insurance balance out. Unemployment benefits are proportional to individual productivity: since wages are a direct function of productivities, our specification boils down to assume that unemployment benefits in the model reflect the potential salary of an household, and the higher the potential salary, the higher the unemployment benefit she is entitled to. In sum:

$$G = \int_{a_{min}}^{a_{max}} b(a)\mathbb{U}(a)da = \int_{a_{min}}^{a_{max}} \tau w(a)\mathbb{E}(a)da = T$$

where $\mathbb{U}(a)$ is the distribution of unemployed households – including singles and married – and $\mathbb{E}(a)$ is the distribution of employed households – including singles and married.

3.3 Marriage

In our economy, agents are born single and can choose between marrying or remaining single at any time during their life-time. Single individuals of any employment status and productivity face the probability to meet potential partners, whose distribution is defined over individual productivities and labor market outcomes. Search in the marriage market is random, and meetings happen with intensity m, a parameter that captures the frequency with which individuals meet on the marriage market. When a meeting happens, agents can choose whether to form a couple or not: a first consideration is that, due to income sharing and concave utility, matching is in principle advantageous for households, as the labor income of a spouse is likely to offer informal insurance against both idiosyncratic and labor market risks, and against subsequent drops in consumption.

However, since being in a couple entails sharing incomes within the household, marriage can be a costly choice for agents, especially when the partner experiences adverse idiosyncratic or labor market shocks. Hence, not all agents prefer matching rather than remaining single, regardless of their productivity level and/or job status, which engineers a mechanism of *selection* of high-productivity individuals into joint-households. Along similar lines, some agents prefer deferring the decision to match with a partner if they expect to have a higher wage in the future. Such incentive may not only lead single households to search longer for a partner, but also to postpone matching in order to first accumulate skills and reach a higher income.[14] Finally, successful matches require both agents in the couple to agree upon forming a joint-household, which we refer to as *bilateral agreement*. Under this assumption, the model naturally produces sorting among individuals in couples, as agents with high productivities and/or good job prospects prefer to match with somebody that has equal or greater productivities and job prospects than their own.

To represent the choice faced by single agents, let's consider any single individual i of productivity $a_{i,t}$. First, she can meet a single individual j of given productivity $a_{j,t}$ and employed in a job with salary $w_{j,t}$, and in this case her choice is whether to stay single or to match according to:

- $\max\{W(a_{i,t}, \tilde{w}_{i,t}); WW(a_{i,t}, \tilde{w}_{i,t}, a_{j,t}, \tilde{w}_{j,t})\}$ if individual i is employed;
- $\max\{U(a_{i,t}); UW(a_{i,t}, a_{j,t}, \tilde{w}_{j,t})\}$ if individual i is unemployed.
- $\max\{O(a_{i,t}); OW(a_{i,t}, a_{j,t}, \tilde{w}_{j,t})\}$ if individual i is outLF.

Secondly, she could meet a possible partner of given productivity $a_{j,t}$ who is unemployed. Her choice in this case is whether to stay single or marry according to:

- $\max\{W(a_{i,t}, \tilde{w}_{i,t}); WU(a_{i,t}, \tilde{w}_{i,t}, a_{j,t})\}$ if individual i is employed;
- $\max\{U(a_{i,t}); UU(a_{i,t}, a_{j,t})\}$ if individual i is unemployed.
- $\max\{O(a_{i,t}); OU(a_{i,t}, a_{j,t})\}$ if individual i is outLF.

Finally, she could meet a possible partner of given productivity $a_{j,t}$ who is out of the labor force. Her choice in this case is whether to stay single or marry according to:

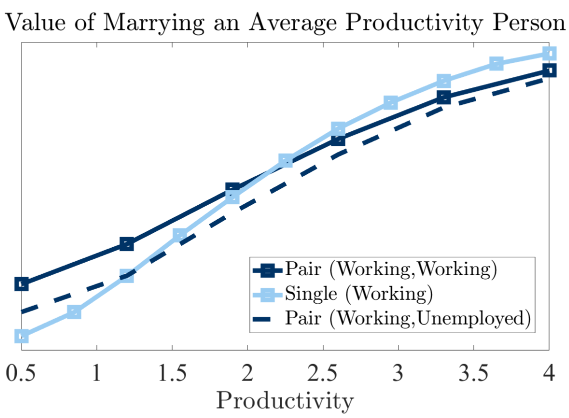
[14]We note that, the concavity of the utility function and the fact that paired agents pull together their respective incomes in an additive way, imply that the payoff function for agents matching on the marriage market is not supermodular. Supermodularity of the payoff functions boils down to the concept of *complementarity* between agents' types. We have instead an environment characterized by *substitutability* between agents' incomes.

- $\max\{W(a_{i,t}, \tilde{w}_{i,t}); WO(a_{i,t}, \tilde{w}_{i,t}, a_{j,t})\}$ if individual i is employed;
- $\max\{U(a_{i,t}); UO(a_{i,t}, a_{j,t})\}$ if individual i is unemployed.
- $\max\{O(a_{i,t}); OO(a_{i,t}, a_{j,t})\}$ if individual i is outLF.

Individual j of productivity $a_{j,t}$ will likewise face a mirrored maximization problem, and a match will be formed only if both agents agree to match. Importantly, we assume here that there is no divorce possibility: once married, households stay married. While our analysis does not rely on this assumption in particular, it is definitely one that we can relax in future work, especially in order to better capture the fact that divorce has gained demographic relevance in the past decade.[15]

To better understand the choice of marrying, Figure 3.1 illustrates the value functions upon matching with a partner characterized by the mean of the productivity distribution in the economy. We consider the cases in which the potential partner is employed or unemployed, and compare them with the option of remaining single. Since the joint surplus is shared within the couple, matching with a person who is employed generally implies a higher value than matching with an unemployed agent, as salaries are higher than unemployment benefits. At the same time, single and employed individuals in the right tail of the productivity distribution have an incentive to search longer for a partner whose productivity (and labor market outcomes) are higher than the average. For this reason, as the value of remaining single is sometimes higher than the one of marrying, the model generates selection and sorting into couples, despite the informal insurance provided within the couple against idiosyncratic and labor market risks.

Figure 3.1: Matching Problem



Children: Agents of any marital status are subject to exogenous fertility. The time of arrival follows a Poisson distribution, where the arrival rate for paired individuals – denoted by π^p – is allowed to be bigger than the Poisson arrival rate for single individuals – denoted by π^s. This assumption is consistent with the fact that married agents are more likely to have children. With respect to the exogenous fertility arrival time t_c we therefore have the following two cases:

$$t_c \sim \begin{cases} Pois(\pi^s), & \text{for single individuals} \\ Pois(\pi^p), & \text{for married individuals} \end{cases} \tag{34}$$

[15] Divorce could represent the risk of loosing the within household's insurance provided by spousal income, which may strengthen ex-ante selection into couples. In that case, our quantified effects would be conservative estimates.

Both π^p and π^s are to be interpreted as the time rates regulating the number of events – i.e. arrivals of a kid – that happen per unit of time t. The probability distribution function of t_c is given by:

$$f(t_c) = \pi^{\mu} e^{-\pi^{\mu} t} \qquad \text{for } \textit{married status } \mu \in \{single; married\} \tag{35}$$

Parents of any marital status $\mu \in \{s; p\}$ are subject to a cost c, interpreted as a per-period cost related to either childcare, or health, education, clothing or other expenses concerning child well-being and upbringing. Fertility costs do not differ between single and married agents in our framework, however, two observations are to be made. On the one hand, married households are more likely to incur the child-related cost c, as $\pi^p > \pi^s$. Yet, differently from singles, they can split such expenses with their partner. On the other hand, when deciding to marry, individuals anticipate that they will incur more likely into fertility costs c. In this sense, even if the previously highlighted mechanisms of selection and sorting in the marriage market hold notwithstanding the inclusion of kids in the model. At the same time, exogenous fertility and, in particular, fertility-related expenses further contribute to the selection of agents into marriage and to sorting between partners.

3.4 Characterization of Agents' Problem

Finally, we characterize the value functions of agents in our economy given their idiosyncratic productivities, labor market outcomes and marital status. To start, take a single employed household characterized by productivity $a_{i,t}$ and working in a job characterized by the wage $w_{i,t}$. Her value function on the labor market is given by the following expression:

$$\begin{aligned}
\rho W(a_{i,t}, \tilde{w}_{i,t}) = {} & \underbrace{u\left((1-\tau)\tilde{w}_{i,t} e^{a_{i,t}}\right)}_{\text{instantaneous utility}} + \underbrace{\delta(U(a_{i,t}) - W(a_{i,t}, \tilde{w}_{i,t}))}_{\text{if she gets unemployed}} \\
& + \underbrace{\xi(0 - W(a_{i,t}, \tilde{w}_{i,t}))}_{\text{if she retires}} + \underbrace{\pi^s(W(a_{i,t}, \tilde{w}_{i,t}; \{kid = 1\}) - W(a_{i,t}, \tilde{w}_{i,t}))}_{\text{if she has a kid}} \\
& + m\Big(\int\int \max\{WW(a_{i,t}, \tilde{w}_{i,t}, a_{j,t}, \tilde{w}_{j,t}) - W(a_{i,t}, \tilde{w}_{i,t}), 0\} * \\
& \underbrace{* \underbrace{\mathbb{1}(WW(a_{j,t}, \tilde{w}_{j,t}, a_{i,t}, \tilde{w}_{i,t}) > W(a_{j,t}, \tilde{w}_{j,t}))}_{\text{and the partner agrees on matching}} e(a_{j,t}, w_{j,t}) dw_{j,t}}_{\text{if she meets any employed person}} \\
& \underbrace{+ \max\{WU(a_{i,t}, \tilde{w}_{i,t}, a_{j,t}) - W(a_{i,t}, \tilde{w}_{i,t}), 0\} \underbrace{\mathbb{1}(UW(a_j, a_{i,t}, \tilde{w}_{i,t}) > U(a_{j,t})) h(a_{j,t}) da_{j,t}}_{\text{and the partner agrees on matching}}}_{\text{if she meets any unemployed person}} \\
& \underbrace{+ \max\{WO(a_{i,t}, \tilde{w}_{i,t}, a_{j,t}) - W(a_{i,t}, \tilde{w}_{i,t}), 0\} \underbrace{\mathbb{1}(OW(a_{j,t}, a_{i,t}, \tilde{w}_{i,t}) > O(a_{j,t})) o(a_{j,t}) da_{j,t}\Big)}_{\text{and the partner agrees on matching}}}_{\text{if she meets any outLF person}} \\
& + \underbrace{\phi_e(\bar{a} - a_{i,t}) \frac{\partial W(a_{i,t}, \tilde{w}_{i,t})}{\partial a_{i,t}} + \frac{\epsilon^2}{2} \frac{\partial^2 W(a_{i,t}, \tilde{w}_{i,t})}{\partial a_{i,t}^2}}_{\text{change due to the stochastic process of } a_{i,t}} + \frac{\partial W(a_{i,t}, \tilde{w}_{i,t})}{\partial t}
\end{aligned}$$

where ρ is the discount factor, δ is the exogenous job destruction rate, ξ is the probability of retiring, π^s is the probability of having a kid for a single agent, and m is the probability of meeting a possible partner on the marriage market. Note that $u((1-\tau)w(a_{i,t}, \tilde{w}_{i,t}))$ is the utility over the labor market income, which depends on the idiosyncratic productivity of the agent $a_{i,t}$ and the wage offer received from her job $\tilde{w}_{i,t}$, and is hence given by $w = \tilde{w}_{i,t}e^{a_{i,t}}$. Moreover, $dF(w)$, $e(a, w)$, $o(a)$ and $h(a)$ represent the distribution of job wages, single employed individuals, single outLF, and single unemployed agents respectively. Finally, the indicator function $\mathbb{1}(\cdot)$ signals whether the potential partner met on the marriage market agrees on matching. The final value function is evaluated subject to the outside-option constraint, that ensures that agent can decide to live the labor force at any moment:

$$W(a_{i,t}, \tilde{w}_{i,t}) \geq O(a_{i,t})$$

Note that the value functions of a single unemployed agent $U(a_{i,t})$ and a single out of the labor force agent $O(a_{i,t})$ are reported and explained in the Appendix, along with the value functions $UU(a_{i,t}, a_{j,t})$, $WU(a_{i,t}, \tilde{w}_{i,t}, a_{j,t})$, $UW(a_{i,t}, a_{j,t}, \tilde{w}_{j,t})$, $WW(a_{i,t}, \tilde{w}_{i,t}, a_{j,t}, \tilde{w}_{j,t})$, $UO(a_{i,t}, a_{j,t})$, $OU(a_{i,t}, a_{j,t})$, $WO(a_{i,t}, \tilde{w}_{i,t}, a_{j,t})$, $OW(a_{i,t}, a_{j,t}, \tilde{w}_{j,t})$, $OO(a_{i,t}, \tilde{w}_{i,t}, a_{j,t}, \tilde{w}_{j,t})$ of married agents.

3.5 Solving for the Equilibrium

We briefly outline the solution algorithm of the model. Let's focus first on single households: the solution of the model is characterized by three distributions $o(e)$, $h(a)$, $e(a, \tilde{w})$ - tracking outLF, unemployed and employed single agents respectively- and value functions $O(a)$, $U(a)$, $W(a, \tilde{w})$ - for outLF, unemployed and employed single individuals respectively. The system of Hamiltonian-Jacobi-Bellman equations (HJB) is written by using instantaneous generating operators:

$$\begin{aligned}
(\rho + \xi)W(a, \tilde{w}) &= u((1-\tau)\tilde{w}e^a) + \mathcal{A}_{WU}U(a) + \mathcal{A}_{WW}W(a, \cdot) + \frac{\partial W(a, \tilde{w})}{\partial t} \\
(\rho + \xi)U(a) &= u(b) + \mathcal{A}_{UU}U(a) + \mathcal{A}_{UW}W(a, \cdot) + \frac{\partial U(a)}{\partial t} \\
(\rho + \xi)O(a) &= u(h)
\end{aligned}$$

Such that: $W(a, \tilde{w}) \geq O(a)$ and $U(a) \geq O(a)$ The flow intensities of any matrix of type $\mathcal{A}$ depend on the distribution $f(\tilde{w})$ and the value functions. When discretized, $\mathcal{A}_{UW}$ is a transition matrix with non-zero entries for transitions from a given $U(a)$ to each $W(a, \tilde{w})$ s.t. $W(a, \tilde{w}) \geq U(a)$ and transition probability equal to $\lambda_u * f(\tilde{w})$. The maximization between labor market value functions and outLF value is solved as a linear complementarity problem. With the Kolmogorov forward equation (KF) we solve for the distributions $e(a, \tilde{w})$ and $h(a)$:

$$\begin{aligned}
\frac{\partial e(a, \tilde{w})}{\partial t} &= \mathcal{A}^*_{UW}h(a) + \mathcal{A}^*_{WW}e(a, w) - (a, \tilde{w}) \\
\frac{\partial h(a)}{\partial t} &= \mathcal{A}^*_{UU}h(a) + \mathcal{A}^*_{WU}e(a, \tilde{w}) - \xi(h(a) - a_0(a))
\end{aligned}$$

where $\mathcal{A}^*$ denotes the adjacent process and is equivalent to $\mathcal{A}'$ in the numerical solution. The distribution of outLF individuals $o(a)$ is treated as an absorbing state and the corresponding row in the matrix $\mathcal{A}$ is substituted with zeros, apart from the diagonal $-\xi$ that captures the probability of retirement. Hence, $o(a)$ can be calculated as a sum of distributions whenever agents endogenously choose the value $O(a)$. To find the solution, we iterate the following steps until convergence:

1. Substitute $\frac{\partial W}{\partial t}$ and $\frac{\partial U}{\partial t}$ in HJB with $\frac{W_{n+1} - W_n}{\Delta}$ and $\frac{U_{n+1} - U_n}{\Delta}$ to get the update on the value functions for the given distributions.

2. Solve for the stationary solution to KF: $\frac{\partial e}{\partial t} = 0$ and $\frac{\partial h}{\partial t} = 0$ s.t. $\int \int e^s(a, \tilde{w}) da d\tilde{w} + \int h(a) da = 1$ to get the update on the households distributions.

A similar case applies in the case of joint-searchers, but the solutions is characterized by nine additional distributions – $hh(a_{i,t}, a_{j,t})$, $eh(a_{i,t}, \tilde{w}_{i,t}, a_{j,t})$, $he(a_{i,t}, a_{j,t}, \tilde{w}_{j,t})$, $ee(a_{i,t}, \tilde{w}_{i,t}, a_{j,t}, \tilde{w}_{j,t})$, $ho(a_{i,t}, a_{j,t})$, $oh(a_{i,t}, a_{j,t})$, $eo(a_{i,t}, \tilde{w}_{i,t}, a_{j,t})$, $oe(a_{i,t}, a_{j,t}, \tilde{w}_{j,t})$, $oo(a_{i,t}, a_{j,t})$ – and value functions – $UU(a_{i,t}, a_{j,t})$, $WU(a_{i,t}, \tilde{w}_{i,t}, a_{j,t})$, $UW(a_{i,t}, a_{j,t}, \tilde{w}_{j,t})$, $WW(a_{i,t}, \tilde{w}_{i,t}, a_{j,t}, \tilde{w}_{j,t})$, $UO(a_{i,t}, a_{j,t})$, $OU(a_{i,t}, a_{j,t})$, $WO(a_{i,t}, \tilde{w}_{i,t}, a_{j,t})$, $OO(a_{i,t}, a_{j,t})$, $OW(a_{i,t}, a_{j,t}, w_j)$ – over all the possible combinations of $a_{i,t}$, $\tilde{w}_{i,t}$, $a_{j,t}$, $\tilde{w}_{j,t}$ for any given pair of married agents i and j.

4 Quantitative Exercise

In this section, we estimate the model on available US data and assess its performance against several untargeted moments. Our quantified framework replicates the observed marital differences in both wages and unemployment rates, as well as the degree of marital sorting and selection documented empirically. We then discuss the interplay of marriage and job search choices, focusing on when and whom agents choose to marry, and if such decisions interact with their labor markets outcomes. Our analysis disentangles the effects of the two main mechanisms at play in the model. First, we illustrate how the insurance provided by spousal incomes impacts agents reservation wage strategies. Secondly, we quantify the contribution of selection into marriage and sorting among couples in affecting the productivity composition of the single and married samples, and in delivering consistent ex-post differences in labor market outcomes by marital status.

4.1 Calibration

We start by discussing the calibration of the model parameters, summarized in Table 2. Our set up features heterogeneous agents that differ in their idiosyncratic productivity a: we draw the initial a from a log-normal distribution $\log \mathcal{N}(\mu_a, \sigma_a)$, normalize $\mu_a = 0$ and set $\sigma_a = 0.215$ to match the volatility of wages upon entering the labor market, as reported in De Nardi et al. (2020). Moreover, we normalize the mass of jobs to 1 and assume that job offers are drawn from a Pareto distribution of parameters (0,1), which delivers a realistically skewed distribution of wage offers.

In the model time is measured in months, and agents in the model are born and start their working life at 25 years old. Of the fifteen parameters to estimate, three are fixed outside the model, for which we choose values common to most works in the literature. In particular, we pick the coefficient of risk aversion $\gamma = 2$, while we set the discount factor $\rho = 0.005$ to match an average annual interest rate of 5% for the US, and the probability of retirement $\xi = 0.002$ to match an average working life of 40 years. We then need to calibrate the parameters governing the fertility process for singles and married agents. In our model, the arrival of a kid is modelled as a Poisson process with probability distribution given by Equation 35. Using our CPS sample, we compute the shares of married and singles with at least one child in the data, which are 0.72 and 0.2 respectively. We then use these shares to back up the arrival rates of kids for singles and married by noting that:

Table 2: Calibration

Parameter	Value	Description
Exogenously Fitted		
γ	2	Coefficient of risk aversion
ρ	0.005	Discount Factor
ξ	0.002	Probability of Retirement
π^s	0.0004	Fertility Rate for Singles
π^p	0.005	Fertility Rate for Married
c	0.2	Children-Related Expenditure Rate
h_{nk}	3	Home Production if No Kids
h_k	13.5	Home Production if Kids
τ	0.025	Tax to Finance Unemployment Insurance
Endogenously Fitted		
ρ_e	0.01	Persistence in Productivity Process (Employed Agents)
ϵ	0.16	Idiosyncratic Volatility in productivity Process
$\bar{a}$	1.27	Productivity Drift
ρ_u	0.02	Persistence in Productivity Process (Unemployed Agents)
δ	0.026	Exogenous Separation Probability
λ_u	20	Job Arrival Poisson Parameter
m	0.0025	Marriage Intensity

$$\int_0^\infty e^{-\frac{t}{480}} \pi^s e^{-\pi^s t} dt = 0.20 \quad \text{for singles} \tag{36}$$

$$\int_0^\infty e^{-\frac{t}{480}} \pi^p e^{-\pi^p t} dt = 0.72 \quad \text{for married} \tag{37}$$

where $\int_0^\infty e^{-\frac{t}{480}} dt$ is the probability of not being retired at time t, considering that the average (working) life of agents in our model is 40 years and that the calibration is done on a monthly basis. Solving the equations delivers the values $\pi^s = 0.0004$ and $\pi^p = 0.005$, which are the arrival rates of the exogenous fertility process for singles and married agents in a month. Moreover, motivated by the fact that, in the US, the average cost of a child between 0 to 18 years old is estimated to be around 20% of the wage of a working person, we set the fertility-related cost $c = 0.2 * \bar{w}$ such that it amounts to 20% of the average wage in our model economy.[16]

To calibrate the values of home production in the presence and absence of children, we match the shares of singles and married out of the labor force, noting that the percentage of married agents out of the labor force is almost 10 times bigger than the percentage of singles. Consistently, our estimation procedure pins down a value of $h_{nk} = 3$ and $h_k = 13.5$. Finally, we need to calibrate the labor income tax rate levied on all employed households to cover the total amount of public insurance provided by the government to unemployed agents. In particular, while unemployment benefit are proportional to individuals' productivities in our model economy, we also ensure that the average unemployment benefit $\bar{b}$ that individuals received is 40% of the wage distribution mean $\bar{w}$.[17] Using the model-implied fraction of employed and unemployed agents in the economy, we

[16] See `www.usda.gov/media/blog/2017/01/13/cost-raising-child`. Such estimated cost does not include college expenses and may vary for rural versus urban areas, from we abstract in the current analysis.

[17] Unemployment benefits vary between 40% and 70% of the wage mean, see Schaal (2017) and Shimer (2005).

hence set τ to be 2.5% according to:

$$\tau = \frac{\bar{b}}{\bar{w}} * \frac{\% \quad unemployed}{\% \quad employed} = 0.4 * \frac{\% \quad unemployed}{\% \quad employed}$$

We then turn our attention to the calibration of the endogenously fitted parameters, for which we target several moments from the data further summarized in Table 3. Following the strategies in Schaal (2017), Shimer (2005) and Nagypal (2007), the exogenous separation probability δ is set to match an average monthly Employment to Unemployment (EU) rate of 1.4% as in Mankart and Oikonomou (2017). To calibrate the search intensity of unemployed workers, we match the average fraction of agents transitioning from Unemployment to Employment (UE) over two consecutive months, which is 25% in US data, as documented in Mankart and Oikonomou (2017).

Table 3: Targeted Moments

	US Data	Model
Employment to Unemployment Rate	0.014	0.014
Unemployment to Employment Rate	0.25	0.27
Average Share of Married Households	0.60	0.60
Wage Progression over Life-Cycle	0.60	0.63
Wage Dispersion	0.71	0.70
Fraction Whose 1-Year Wage Change < 20%	0.70	0.72
Wage Loss Upon 1 Additional Month Unemployed	0.10	0.09

Focusing on the marriage market, we set the (monthly) marriage intensity $m = 0.0025$ to match the average fraction of married people in CPS data over the last two decades, which is around 60% (see Doepke and Tertilt (2016)). The meeting probability governs the rate at which individuals in the model meet potential partners on the marriage market. As a concrete example of the magnitude of the estimated marriage intensity m, in every year of her life a given agent in our baseline economy has roughly 3 in 100 chances to meet a possible partner. Importantly, this is the only moment related to household's formation that we directly target, whereas the model alone consequently generates sorting and selection into couples, as it will be explained in the next section.

Finally, the volatility of the idiosyncratic shock ϵ in the productivity process of Equation 33 is set to match the average wage variance in the data, as in Kaplan et al. (2018). The productivity drift $\bar{a}$ is calibrated to deliver a realistic wage growth over the life-cycle: in particular, we match the ratio between the average wage and the average wage of first-time employed individuals, as reported in De Nardi et al. (2020). We then turn to the persistence parameters that influence the dynamics of productivity accumulation and depletion for employed and unemployed workers. For ρ_e, we follow Kaplan et al. (2018) and target the fraction of agents that changes their wage by less than 20% in a model period. To calibrate ρ_u, we target instead an average loss of 1% in starting wages after an additional month in unemployment – computed as the percentage change with respect to the wages agents had before the unemployment spell – as documented by Neal (1995).[18]

In the estimation procedure, we minimize the weighted distance between the moments in the data and in the model. For arbitrary values of the vector of parameters to be estimated, we first solve

[18] It is important to note that, similarly to other works in this literature, the persistence parameter ρ_u governing the productivity depletion of unemployed individuals is such that $\rho_u = 2 * \rho_e$. This precisely reflects the fact that productivities are faster lost than accumulated, as argued and documented in Jarosch (2021). In a different setting, Huckfeldt (2016) also finds that the probability of loosing skills is more than twice the one of accumulating skills.

for the equilibrium and evaluate the stationary distributions discussed in Section 3. Using these distributions, we compute the average wage, unemployment benefit and fertility cost, along with the marriage rate, such that they match their data counterparts. Then, we compute the EU and UE transition rates using the transition matrices described at the end of Section 3. Denoting the simulated moments by $\Omega(X)$ and those computed from the data as $\hat{\Omega}$, we estimate the fitted parameters $\hat{X}$ using a minimum distance criterion given by:

$$L(X) = \min_{X}(\hat{\Omega} - \Omega(X))' W(\hat{\Omega} - \Omega(X))$$

We set the weighting matrix W = I and use grid search to find the minimum.

4.1.1 Untargeted Moments

We validate our framework by assessing how it performs on untargeted dimensions related to the labor market outcomes of single and married agents, as well as to US marital patterns. First, the model can replicate the observed differences in the average salaries of married and single individuals: the ratio of their wage means moderately overfits its data counterpart, but matches 75% of the WMP and improves the fit of existing models by 25%. Two mechanisms determine this result: on the one hand, intra-household insurance enables partners to wait in unemployment longer and choose better paid jobs. On the other hand, by introducing endogenous marriage in the model, we also allow individuals to select and sort into the married pool, which, ex-post, is composed by households of higher productivity and better labor market outcomes. The composition of the married sample therefore contributes to the observed higher wage mean for agents in couples. Through the lens of our model, the WMP is the result of both a (i) process of selection-into-marriage, and a (ii) higher reservation wage strategy of individuals in joint-households.

The second validation of our quantitative framework is with respect to the UMG. In the absence of endogenous household's formation, the typical mechanism at work in joint-search models is that married individuals wait in unemployment longer before accepting a job offer, resulting in higher average unemployment rates. However, as shown in Section 3, unemployment rates are lower for agents in couples. Endogenizing marriage leads to heterogeneities in the productivity composition of single and joint-households that ensure a correct prediction for the unemployment rate differences between singles and married. Due to selection effects, married agents are on average of higher productivities and better job prospects, and hence have lower unemployment rates. In particular, the model predicts that the unemployment rate of married individuals is 1.3 p.p. lower than the one of singles, and hence explains 50% of the difference observed in US data.

Focusing instead on marital patterns, the model replicates the increasing share of married over individual productivity,[19] and matches 80% of the elasticity of marriage rates to household's skills, which is consistent with the phenomenon of selection into couples observed empirically. There are two additional observations on the timing and outcome of the marriage choices in our set up. First, individuals wait to find a suitable partner and, due to search frictions in the marriage market, such process takes time. Accordingly, the average age at marriage is 31 years old in the model,[20] close to the one reported for the US in the past years. Secondly, agents prefer to secure a good wage before forming a joint-household, as marriage can be costly in case of adverse labor market

[19] As previously discussed, what in the model constitutes an agent's type is her initial level of productivity and progression. For the sake of the comparison in this section, we proxy the *productivity* level in our model with the *educational attainment* of households in the data. This is an imperfect mapping, but does not affect our main message.

[20] Recall the agents in the model are considered since they enter the labor market at 25 years old.

Table 4: Main Untargeted Moments

	Data	Model
Labor Market		
Unemployment Marital Gap	2.5 pp	1.3 pp
Wage Marital Premium	0.20	0.25
Marriage		
Average Marital Sorting	0.52	0.35
Elasticity of Share of Married and productivity	0.19	0.22
Average Age at Marriage	29yo	31yo

outcomes of the spouse, with whom the income is shared. This mechanism is reinforced by the fact that the model allows for exogenous fertility: when choosing to marry, agents also discount in their decision the fact joint-households are more likely subject to fertility-related costs.

In addition to that, costly marriage and bilateral agreement upon matching allow the model to replicate 67% of the degree of sorting among couples. The fact that agents tend to marry "their like" is an empirical regularity that has been extensively documented in US data (see Greenwood et al. (2014) for example), where sorting can relate to the positive correlation of wealth profiles or personal characteristics, such as education, between individuals in couples. We abstract from wealth dynamics and instead focus on educational attainments, which proxy individuals' productivities in the model. Accordingly, we measure sorting in CPS as the correlation between married agents' educational attainments, while in the model, we compute the correlation in partners' productivities upon matching and find that it slightly under-predicts the empirical one.

Table 5: Couples' Shares in Working Age Population

	(E,E)	(E,O)∪(O,E)	(E,U)∪(U,E)	(U,U)	(U,O)∪(O,U)	(O,O)
Model	0.573	0.278	0.062	0.001	0.054	0.037
Data	0.606	0.324	0.042	0.004	0.012	0.014

Notes: Empirical shares of couples in the working age population in each labor market status computed from Mankart and Oikonomou (2016) excluding retired individuals. Unemployed, employed and outLF individuals are 85% of the population.

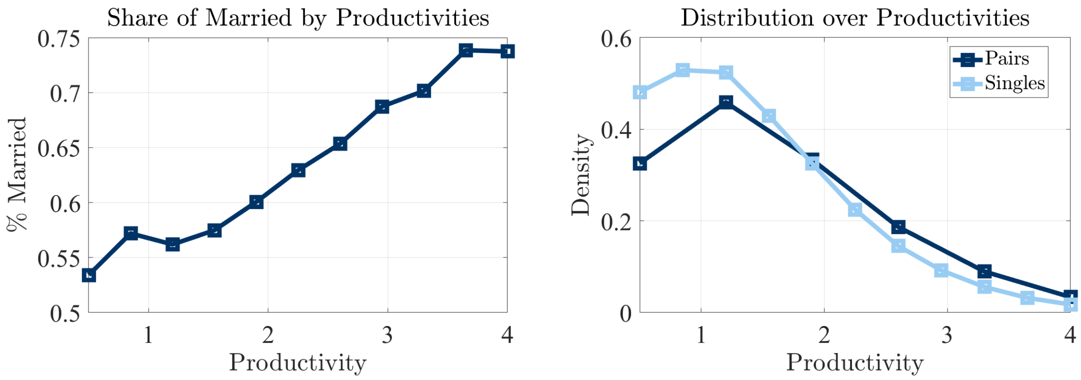
Finally, Table 5 compares the shares of couples in different labor market statuses in the model and in the data. In particular, we take moments from Mankart and Oikonomou (2017), who compute the shares of employed-employed (EE), employed-outLF (EO), employed-unemployed (EU), unemployed-unemployed (UU), unemployed-outLF (UO) and outLF-outLF (OO) couples in the population using CPS. Since Mankart and Oikonomou (2017) consider also retirees in their estimation, we then weight their computed shares by the percentage of agents into the working age population for a better comparison with our model-implied counterparts. Our framework can satisfactorily match *endogenously* the shares of couples in different labor market statuses, but slightly over-predicts the share of UO and OO couples and under-predicts the share of UU ones.

In conclusion, by endogenizing the decision to marry into a joint-search framework, it is possible to reconcile both wage and unemployment differences by marital status, and to also reproduce marital patterns consistent with empirical observations. As such, the informal insurance against idiosyncratic and labor market shocks provided within couples is certainly a key channel to consider in joint-search models, but its power in explaining labor market differences by marital status is nonetheless limited to the WMP. Once selection into joint-households is accounted for, the model can improve the quantitative fit of the WMP, and can finally explain 2/3 of the observed UMG.

4.1.2 A Further Analysis of the Marriage Market

The key novel feature of our framework is to embed the decision of forming a joint-household in an otherwise standard joint-search set up. In particular, the incentive to match is enhanced by the fact that agents have concave utility over consumption and cannot self-insure against idiosyncratic and unemployment risks by saving.[21] However, we have argued before that matching with someone can be costly in some states of the world, particularly when a spouse is hit by adverse idiosyncratic or labor market shocks and precisely because income is shared within the couple. As such, since marrying requires bilateral agreement, single agents in the economy will show a different propensity to marry (and sorting) based on their productivities and labor market status.

Figure 4.1: Marriage Rates and Distribution over Productivities



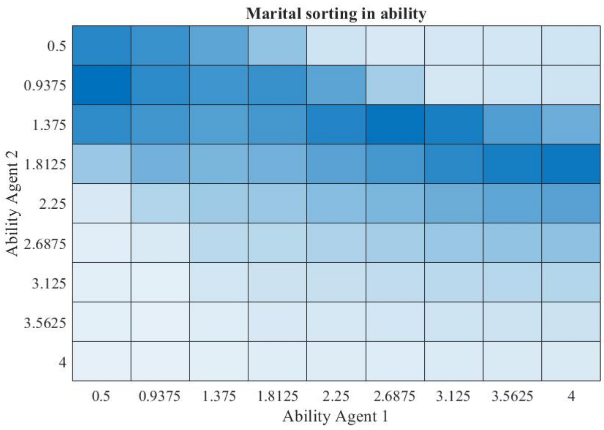
In the left panel of Figure 4.1, we plot the share of married agents by productivity level. Individuals with higher productivities have higher marriage rates, as they tend to have on average better paid jobs and be more likely to afford the costs associated to form and maintain a joint-household. As further confirmed by the right panel of Figure 4.1, which plots the distribution of agents over productivity levels, married individuals also have a higher mean productivity. This comparison illustrates the compositional heterogeneities in the sample of married and singles households, in that agents select into marriage based on their productivity and job prospects.[22]

As reported in Table 4, our framework delivers a positive correlation between the productivities of individuals in couples upon matching, which fits roughly 70% of its empirical counterpart computed using CPS data for the US. In Figure 4.2, we further plot the probabilities of marrying for agents of different productivities. Darker colors are associated with a higher likelihood of matching, which means that individuals tend to positively sort along their productivity levels and

[21] Abstracting from savings is a simplifying assumption that keeps the model tractable. As a shorthand for the different self-insurance across individuals of heterogeneous productivities, we have assumed that the unemployment compensation is increasing in productivity, which hence ensures highly productive agents a greater consumption level when in unemployment. If agents in the model had been able to self-insure against unemployment by saving, the consumption of highly productive agents when unemployed would have been similarly higher.

[22] We would obtain the same prediction if we were to focus on the distribution of singles and married over wages.

Figure 4.2: Marital Sorting



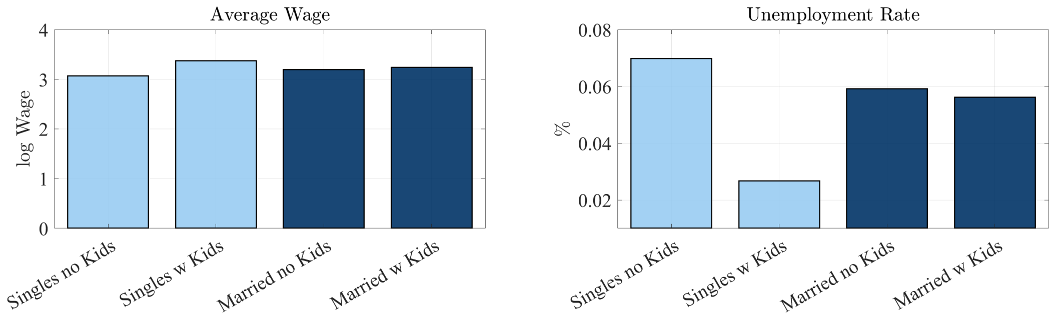
are more likely to marry partners of similar labor market outcomes. Importantly, households have an incentive to marry, but they do not match with every single agent in the distribution of possible partners. Bilateral agreement and costly household's formation push agents to match with partners of equal or better productivity and employment outcomes. Moreover, search frictions in the marriage market reinforce marital sorting, as single agents do not meet every period, but, at the same time, they do expect to be meeting other singles with some probability during their life-time. This feature allows single individuals to postpone matching and keep searching for a partner.

4.1.3 Analysing Labor Market's Outcomes

We now take a closer look at the labor market performance of agents in our economy. In Figure 4.3, we report the average wage (left panel) and unemployment rate (right panel) of married and singles. Due to the compositional differences across the single and married pools of individuals, agents in couples have higher productivities, and high-productivity workers in turn have better labor market outcomes. Ex-post, this contributes to explain why married agents have higher wages and lower unemployment rates compared to singles. Similarly, one can focus on individuals with and without kids: since exogenous fertility in the model entails children-related expenses for households, individuals with kids tend to have on average lower unemployment rates and wages compared to those without kids, regardless of their marital status. Exogenous fertility pushes unemployed agents to accept relatively worse job offers and quit unemployment faster to sustain children-related expenses without incurring in substantial consumption drops. This effect is exacerbated by the fact that individuals cannot save to smooth their consumption profile.

The model also allows to disentangle the relative contribution of within household's insurance and selection into couples to the differential labor market outcomes by marital status, as further illustrated in Table 6. Since the mechanism of informal insurance provided within households alone would imply higher unemployment rates for married individuals, such channel cannot deliver predictions consistent with the UMG documented in the data. A set up without endogenous marriage has 0 fit on the unemployment gap between singles and married, predicting a -2 p.p. gap instead of the documented +2.5 p.p. gap. When endogenous marriage is included in the model, single households have 1.3 p.p. higher unemployment rates compared to married, which fits 50% of its empirical counterpart. In this sense, by allowing agents to endogenously marry, select, and sort into couples, the model is able to imply a higher share of married agents among

Figure 4.3: Average Wages and Unemployment Rates



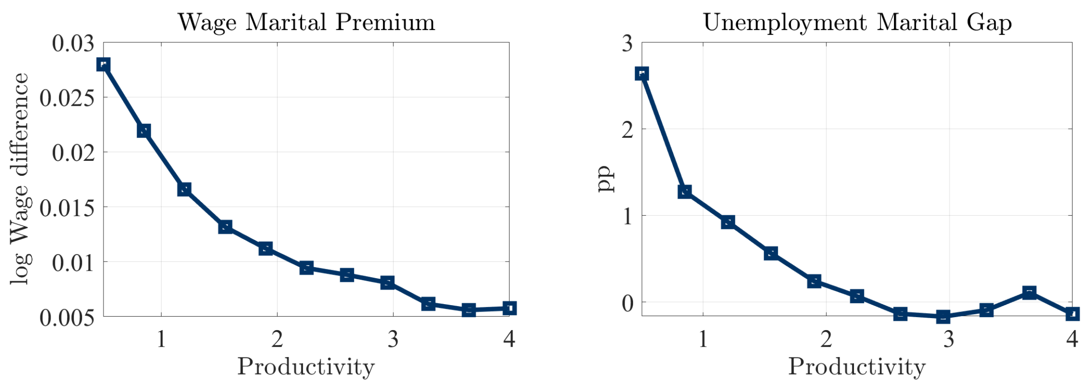
the high-productivity and low-unemployment households, and reconciles the theory with the data.

Table 6: Decomposition

	Wage Marital Premium	Unemployment Marital Gap
data	0.20	2.5 pp
model without endogenous marriage	0.12	-2.0 pp
model with endogenous marriage	0.25	1.3 pp

However, both within household's insurance and selection into couples contribute to higher wages for married individuals. The WMP is due to an interplay between the compositional differences in productivities across the married and single samples and the fact that unemployed married job searchers can wait longer for a better wage offer. If we were to shut down the endogenous marriage decision, the quantitative fit of the model in terms of WMP will be lower, replicating only 60% of the wage differences between married and singles (consistent with results from Pilossoph and Wee (2020)). On the contrary, allowing for endogenous marriage leads the model to fit 75% of the WMP, which represents a 25% improvement with respect to previous studies on joint-search.

Figure 4.4: Wage and Unemployment Spells over Productivities



As an additional remark, Figure 4.4 plots wage and unemployment rates differences across single and married agents over their productivity. When the composition of the married and single sample is controlled for, marital differences in wages and unemployment rates become small and less significant. In particular, the within household's insurance mechanism that allows married agents to have a higher reservation wage, is left to explain a smaller fraction of the log wage difference between married and singles, as depicted in the left panel of Figure 4.4. Focusing on the right panel, low-productivity married agents have still lower unemployment rates compared to singles. This is due to the fact that married households tend to have kids, which represents an incentive to quit unemployment quickly, especially for low-productivity individuals. We then conclude that, if we were able to equally control for the actual extent of selection into couples also in the data, we should find less stark and significant marital differences in wages and unemployment rates.

Figure 4.5: Wage and Job Changes After a Period of Unemployment

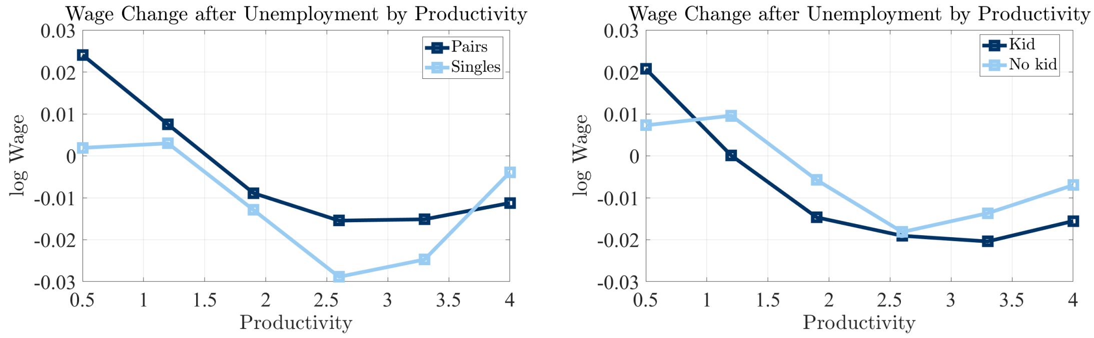


Figure 4.5 finally shows that, upon re-employment, agents tend to experience a drop in their salary compared to their pre-unemployment earnings. On the one hand, high-salary jobs are relatively scarce and take long to find due to the skewed profile of the wage offer distribution. On the other hand, agents deplete their productivities while unemployed, which negatively affects their efficiency-units wage upon re-employment. Controlling for the productivity composition of the singles and pairs' sample, the performance of single individuals upon re-employment is better. This is due to the fact that the presence of children-related expenses in the household's budget constraint constitutes an incentive to quit unemployment faster and accept relatively worse job offers, as documented in the right panel of Figure 4.5. Since fertility rates are higher for couples, married households experience larger average drops in wages after a period in unemployment.

5 Optimal Unemployment Benefits

In this last section, we explore the appropriateness of public policies aimed at improving the welfare of both single and joint-households. We first note that, in our model economy, labor force participants are subject to uninsurable idiosyncratic productivity shocks, while employed individuals face also the risk of becoming unemployed. Both circumstances decrease the disposable income of agents, and, in particular, unemployment benefits replace only a fraction of the household's salary. In addition to that, we have allowed for exogenous fertility for both singles and married households in the model: kids are costly for agents, who have to pay about 20% of the their average income in children-related expenses each period. This cost represents a non-negligible item in households' budget constraint, particularly when they suffer adverse labor market shocks.

Yet, even if the volatility of productivity shocks and the job destruction rate are the same across individuals, couples suffer less from idiosyncratic and unemployment shocks because incomes are pulled together within joint-households. As an example, unemployed workers paired to partners that are employed may enjoy higher level of consumption as opposed to single unemployed workers. On the other hand, even if the cost of raising a kid in our model economy is in principle the same across households, and married individuals are more likely to be subject to exogenous fertility, children-related expenses can be shared among agents in joint-households, while single individuals have to face them alone. Given these considerations, we aim to assess whether the fiscal surplus obtained from the labor income tax collection could be re-distributed towards households with less insurance against labor market risks, especially if they face children-related expenses.

We begin by recalling that, in our baseline economy, labor income taxes are levied by the government on all working individuals and used to finance unemployment benefits. In particular, we denote by G the total fiscal surplus collected by the public sector on all workers, regardless of their marital status, the labor market status of their spouse (if the agent is married), and the presence of kids in the household, which is given by:

$$G \equiv \int_{(a,w)\in E} \tau w e^{a} d(a,w)$$

where E is the sum of the distributions of working agents, given by:

$$E \equiv \Big(e(a;\{kid \in \{0,1\}\}) + ee(a_i, a_j;\{kid \in \{0,1\}\}) + eh(a_i, a_j;\{kid \in \{0,1\} + eo(a_i, a_j;\{kid \in \{0,1\}\} \Big)$$

Total fiscal surplus G is redistributed to cover total unemployment benefits B, defined as follows:

$$B \equiv \int_{a\in H} b(a) da$$

where H is the sum of the distributions of unemployed agents, given by:

$$H \equiv \Big(h(a;\{kid \in \{0,1\}\}) + he(a_i, a_j;\{kid \in \{0,1\}\}) + hh(a_i, a_j;\{kid \in \{0,1\} + ho(a_i, a_j;\{kid \in \{0,1\}\} \Big)$$

where $b(a)$ is the progressive individual unemployment benefit paid to unemployed agents, regardless of their marital and fertility status. The spirit of the following exercises will be to keep the labor income tax rate τ fixed and instead understand how to optimally distribute the total fiscal surplus G according to different criteria. For example, we analyze if unemployment benefits could be designed to depend on the family type of individuals, and assess whether publicly-provided insurance should entail a higher compensation for single agents with no access to within household's insurance. Parallel to that, we ask if part of the implied fiscal surplus from taxes could be disposed in the form of a child credit for both working and non-working parents.

Importantly, we want to find the optimal *combination* of these two policies, but, for the sake of the exposition, we first proceed to explore the appropriateness of a redistribution of publicly collected labor income taxes in favor of households with kids, both single and married, as they are the ones subject to children-related expenses that decrease their disposable income.[23] One can find examples of policies in favor of families with children not only in the US but also in many other developed and developing countries. Such policies can be classified in different broadly-defined categories: (i) subsidies, which may take the form of health and educational costs assistance programs, housing vouchers or food stamps; (ii) services, such as publicly-provided childcare; or (iii) tax credits (for an extensive discussion on child-related transfers in the US, see Guner et al. (2020)).

While it is beyond our scope to provide an extensive investigation of child-related programs, we

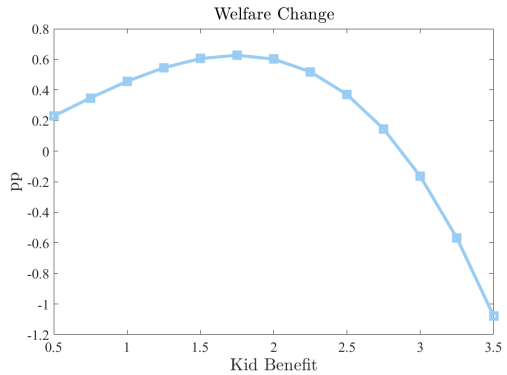
[23] Similarly in Guner et al. (2020), who study family policies but not their interplay with unemployment benefits.

interpret our exercise as similar in spirit to allowing for child tax credits in our model economy. This is motivated by the fact that the source of fiscal resources rebated in favor of families are labor income taxes. Moreover, we want to consider a scenario in which *all* labor force participants, both unemployed and employed, can receive such credit, differently from the situation in which *only* workers can be eligible for child benefits, as it is the case for most of the current US child-related policies. To that end, we modify the government budget constraint to read as follows:

$$
\begin{aligned}
G = B - T \times \mathbb{U} + \\
+ \left(\frac{\theta_k}{2}\right) \Big(h(a;\{kid=1\}) + he(a_i,a_j;\{kid=1\}) + hh(a_i,a_j;\{kid=1\}) + ho(a_i,a_j;\{kid=1\}) \Big) - \\
- \left(\frac{\theta_k}{2}\right) \Big(h(a;\{kid=0\}) + he(a_i,a_j;\{kid=0\}) + hh(a_i,a_j;\{kid=0\}) + ho(a_i,a_j;\{kid=0\}) \Big)
\end{aligned}
$$

where B are the unemployment benefits, θ_k is a kid specific term which can be either positive, if resources are shifted towards agents with kids, or negative, if resources are redistributed in favor of individuals without kids, and T is a residual term to balance the government budget via lump-sum transfers (note that $\theta_k = 0$ and $T = 0$ in the baseline economy). The optimal policy is the one that maximizes aggregate welfare, computed as the sum of utilities over income weighted by marginal utilities. As shown in Figure 5.1, it is beneficial to redistribute approximately 2 income units from the fiscal surplus of the economy in favor of parents, regardless of their marital status.[24]

Figure 5.1: Welfare Changes Under Child Benefits



We add to this analysis by combining this first margin of fiscal redistribution together with the provision of unemployment benefits. Keeping the labor income tax unchanged, we further allow the unemployment benefit administered by the government to depend on the marital status of the recipient. In particular, this second policy dimension explores if it is beneficial to redistribute resources from agents in couples to bachelor individuals, who have no access to within household's insurance against both idiosyncratic and unemployment risk. In fact, single unemployed agents have on average a higher marginal utility of consumption, being those whose fall in disposable income when unemployed is harsher, as also documented in Choi and Valladares-Esteban (2019). To assess both margins of fiscal redistribution together and to consider both the presence of kids and the marital status of the households in the optimal provision of unemployment benefits, we

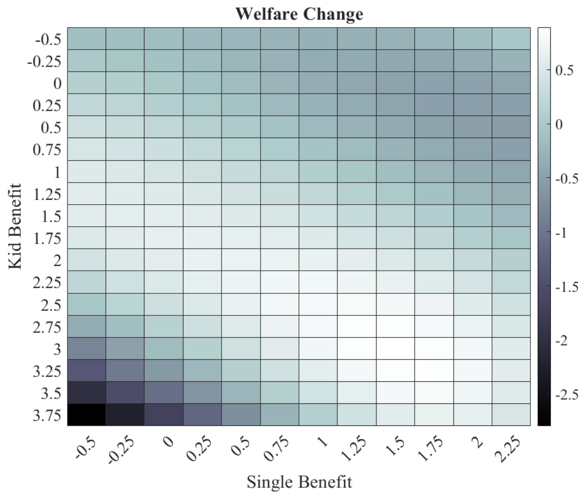
[24] An important caveat is that, in our model, fertility is exogenous, consistent with the modeling of fertility arrival in Guner et al. (2020). Arguably, fertility decision are mostly and endogenous choice for individuals: when we hence refer to such cost c we have in mind the cost component of raising children, being the expenses expected or unexpected. If the cost component c itself was the result of a choice variable, our normative analysis would be partially different.

then modify the government budget constraint to read as follows:

$$G = B - T \times \mathbb{U} + \left(-\frac{\theta_k}{2} + \frac{\theta_b}{2}\right)\Big(h(a; \{kid \in \{0\}\})\Big) + \left(\frac{\theta_k}{2} + \frac{\theta_b}{2}\right)\Big(h(a; \{kid \in \{0\}\})\Big) +$$
$$+ \left(\frac{\theta_k}{2} - \frac{\theta_b}{2}\right)\Big(he(a_i, a_j; \{kid \in \{1\}\}) + hh(a_i, a_j; \{kid \in \{1\}\}) + ho(a_i, a_j; \{kid \in \{1\}\}\Big) +$$
$$+ \left(-\frac{\theta_k}{2} - \frac{\theta_b}{2}\right)\Big(he(a_i, a_j; \{kid \in \{0\}\}) + hh(a_i, a_j; \{kid \in \{0\}\}) + ho(a_i, a_j; \{kid \in \{0\}\}\Big)$$

where B are the unemployment benefits, θ_k is the child-specific term described before, and θ_b is a specific term which can be either negative, if resources are shifted towards unemployed individuals in couples, or positive, if resources are instead redistributed in favor of single unemployed agents. Again, T is simply a residual term to balance the government budget via lump-sum transfers (note that $\theta_k = 0$, $\theta_b = 0$ and $T = 0$ in the baseline economy). We hence search for the combination of θ_k and θ_b that achieves the highest aggregate welfare in our model economy.

Figure 5.2: Optimal Policy

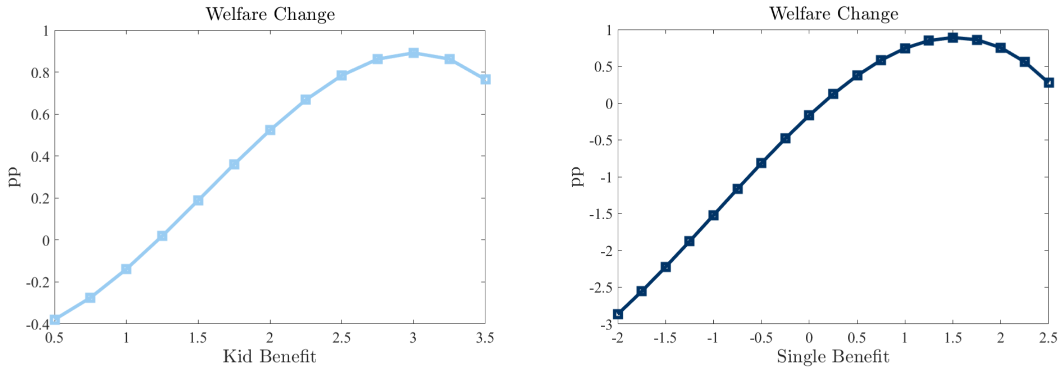


The results are shown in Figure 5.2: in the picture, lighter colors correspond to the highest welfare gains, as opposed to darker colors. The optimal policy in this economy is one that acts along both margins considered to better insure agents against income falls and additional costs in which they may incur during their life-time, and it is able to generate a total of 0.89% welfare gain for the economy as a whole. In conclusion, we stress again that, on the one hand, it is beneficial to redistribute from joint to single households, who are less insured against drops in consumption following an unemployment shock: as such, it is optimal to increase the unemployment benefit of single individuals by 1.5 income units. On the other hand, the optimal policy also implies a subsidy of 3 income units in favor of any household with a kid. These two dimensions of redistribution are further shown on separate panels in Figure 5.3.

6 Conclusion

In this paper, we have presented and quantified a model of joint-search in the labor markets characterized by three novel elements: endogenous household's formation, exogenous fertility, and heterogeneity on the worker's side. In particular, by endogenizing the marriage decision, we were

Figure 5.3: Welfare Changes Under Different Policies



able to fit the differences in both the average wage and the unemployment rate between singles and married households observed in US data. Prior research in this regard had been successful in explaining the WMP only, at the expense of generating counterfactual predictions on the unemployment rates' differences by marital status. In this sense, we see our project as a way to extend the workhorse model of joint-search à la Guner et al. (2012) to allow for a further analysis of the heterogeneities in labor market outcomes across individuals of different marital status.

Empirically, married agents report higher wages but lower unemployment rates. In our model economy, agents select and sort into joint-households, which induces compositional differences in the samples of singles and married that are key in explaining the WMP and the UMG. The insurance provided within the household against both idiosyncratic and labor market risks, contributes to the quantitative fit of the WMP, whereas its explanatory power with respect to the UMG is null. When endogenous marriage is allowed, the model can instead fit 75% of the WMP and 65% of the UMG. Properly calibrated on US data, our model can also replicate well several untargeted moments of the marriage market, such as the positive assortative matching of agents in couples, and the increasing profile of marriage rates over individuals' skills (proxied by educational attainment in the data). Finally, we have used our framework as a laboratory to study fiscal redistribution across agents, and shown that the optimal policy combines higher benefits for single households and for any household that faces child-related costs, irrespective of their marital status.

Appendix

A Empirical Analysis

A.1 Wages and Unemployment

In the following section, we unpack the results reported in Section 2 by analysing the regression outcomes of Equation 32. We first document the differences in the hourly and yearly wages of single and married individuals and then consider their unemployment rates. Throughout the analysis, we make use of the CPS sample discussed in the main text, and further expand the period to include the years between 1990 and 2019 (hourly wages are available in CPS starting from 1990).

Table A1: Wages in CPS

	(log) Hourly Wage	(log) Yearly Wage	(log) Hourly Wage	(log) Yearly Wage
Married	0.2228***	0.4093***	0.1026***	0.1828***
	(0.0054)	(0.0108)	(0.0088)	(0.0177)
Controls	N	N	Y	Y
Year FE	N	N	Y	Y
Occupation FE	N	N	Y	Y
Observations	44,950	435,186	20,089	195,464
R^2	0.0472	0.0039	0.4155	0.2820

Notes: Robust standard errors in parentheses. Survey weights are used. ***p<0.01, **p<0.05, *p<0.1. Control variables include gender, education, employment status in the previous calendar year, ethnicity, number of kids, age, spousal education and spousal current employment status.

As reported in Table A1, whether we consider the hourly wage or the yearly one, being married is associated with higher labor income, and the result is robust to the inclusion of several controls and year and occupation fixed effects. In Table A2, we instead show that being married is also associated with a lower likelihood of being unemployed (both with and without controls).

Table A2: Unemployment in CPS

	Unemployment	Unemployment
Married	-0.0418***	-0.0251***
	(0.0012)	(0.0018)
Controls	N	Y
Year FE	N	Y
Previous Year Occupation FE	N	Y
Observations	393,309	176,716
R^2	0.0065	0.1415

Notes: Robust standard errors in parentheses. Survey weights are used. ***p<0.01, **p<0.05, *p<0.1. Control variables include gender, education, employment status in the previous calendar year, ethnicity, number of kids, age, spousal education and spousal current employment status.

B Model

B.1 Value Functions

In this section, we report all the relevant value functions of the model, for both single and married agents, whether they are unemployed, employed, or out of the labor force. In the main body of the paper, we have discussed the value function of a single working individual $W(a_i, \tilde{w}_i)$; here we give details on the following value functions: $UU(a_i, a_j)$, $WU(a_i, \tilde{w}_i, a_j)$, $UW(a_i, a_j, \tilde{w}_j)$, $WW(a_i, \tilde{w}_i, a_j, \tilde{w}_j)$, $UO(a_i, a_j)$, $OU(a_i, a_j)$, $WO(a_i, \tilde{w}_i, a_j)$, $OW(a_i, a_j, \tilde{w}_j)$, $OO(a_i, \tilde{w}_i, a_j, \tilde{w}_j)$. Let's start by the value function of a single unemployed agent, which is given by:

$$
\begin{aligned}
\beta U(a_i) = \; & \underbrace{u(b(a_i))}_{\text{instantaneous utility}} + \underbrace{\lambda_u \int \max\{W(a_i, \tilde{w}_i) - U(a_i), 0\} dF(\tilde{w}_i)}_{\text{if she finds job}} \\
& + \underbrace{\xi(0 - U(a_i))}_{\text{if she retires}} + \underbrace{\pi^s(U(a_i; \{kid = 1\}) - U(a_i))}_{\text{if she has a kid}} \\
& + \underbrace{m\Big(\int\int \max\{UW(a_i, a_j, \tilde{w}_j) - U(a_i), 0\} \text{x} \underbrace{\mathbb{1}(WU(a_j, \tilde{w}_j, a_i) > W(a_j, \tilde{w}_j))}_{\text{and the partner agrees on matching}} e(a_j, \tilde{w}_j) d\tilde{w}_j}_{\text{if she meets any employed person}} \\
& + \underbrace{\max\{UU(a_i, a_j) - U(a_i), 0\} \underbrace{\mathbb{1}(UU(a_j, a_i) > U(a_j)) h(a_j) da_j\}}_{\text{and the partner agrees on matching}}}_{\text{if she meets any unemployed person}} \\
& + \underbrace{\max\{UO(a_i, a_j) - U(a_i), 0\} \underbrace{\mathbb{1}(OU(a_j, a_i) > O(a_j)) o(a_j) da_j}_{\text{and the partner agrees on matching}}\Big)}_{\text{if she meets any outLF person}}
\end{aligned}
$$

where β is the discount factor, λ_u is the probability to find a job and exit unemployment, ξ is the probability of retiring, π^s is the probability of having a kid for a single agent, and m is the probability of meeting a possible partner on the marriage market. Note that $u(b(a_i))$ is the utility over the unemployment benefit. The unemployment benefit is given by $b(a_i) = b * e^{a_i}$ and depends on the idiosyncratic productivity of the agent a_i and the proportional term b set by the government, which is pinned down numerical in the calibration exercise. Moreover, $dF(\tilde{w})$, $e(a, w)$, $o(a)$ and $h(a)$ represent the distribution of job wages, single employed individuals, single outLF, and single unemployed agents respectively. Finally, the indicator $\mathbb{1}$ signals whether the potential partner met on the marriage market agrees on matching.

Moreover, the value function of a single outLF agent is given by:

$$\beta O(a_i) = \underbrace{u(h)}_{\text{instantaneous utility}} + \underbrace{\xi(0 - O(a_i))}_{\text{if she retires}} + \underbrace{\pi^s(O(a_i;\{kid=1\}) - O(a_i))}_{\text{if she has a kid}}$$

$$+ \underbrace{m\Big(\int\int \max\{OW(a_i,a_j,\tilde{w}_j) - O(a_i),0\}\text{x}\,\underbrace{\mathbb{1}(WO(a_j,\tilde{w}_j,a_i) > W(a_j,\tilde{w}_j))}_{\text{and the partner agrees on matching}}e(a_j,\tilde{w}_j)d\tilde{w}_j}_{\text{if she meets any employed person}}$$

$$+ \underbrace{\max\{OU(a_i,a_j) - O(a_i),0\}\,\underbrace{\mathbb{1}(UO(a_j,a_i) > U(a_j))h(a_j)da_j\}}_{\text{and the partner agrees on matching}}}_{\text{if she meets any unemployed person}}$$

$$+ \underbrace{\max\{OO(a_i,a_j) - O(a_i),0\}\,\underbrace{\mathbb{1}(OO(a_j,a_i) > O(a_j))o(a_j)da_j\Big)}_{\text{and the partner agrees on matching}}}_{\text{if she meets any outLF person}}$$

where β is the discount factor, ξ is the probability of retiring, π^s is the probability of having a kid for a single agent, and m is the probability of meeting a possible partner on the marriage market. Note that $u(h)$ is the utility over home production. Moreover, $e(a,w)$, $o(a)$ and $h(a)$ represent the distribution of single employed individuals, single outLF, and single unemployed agents respectively. Finally, the indicator $\mathbb{1}$ signals whether the potential partner met on the marriage market agrees on matching.

We now turn to the value function of married agents in our model economy. Since we do not allow for divorce and remarriage, married agents are not active on the marriage markets, but they still make choices in the labor markets jointly, and are subject to both exogenous fertility and retirement. First, the value function of a couple with both agents employed is as follows:

$$\beta WW(a_i,\tilde{w}_i,a_j,\tilde{w}_j) = \underbrace{u\left((1-\tau)\frac{w(a_i)+w(a_j)}{2}\right)}_{\text{instantaneous utility}} + \underbrace{\delta(UW(a_i,a_j,\tilde{w}_j) - WW(a_i,\tilde{w}_i,a_j,\tilde{w}_j))}_{\text{if she gets unemployed}}$$

$$+ \underbrace{\delta(WU(a_i,\tilde{w}_i,a_j) - WW(a_i,\tilde{w}_i,a_j,\tilde{w}_j))}_{\text{if spouse gets unemployed}}$$

$$+ \underbrace{\xi(0 - WW(a_i,\tilde{w}_i,a_j,\tilde{w}_j))}_{\text{if they retire}}$$

$$+ \underbrace{\pi^p(WW(a_i,\tilde{w}_i,a_j,\tilde{w}_j;\{kid=1\}) - WW(a_i,\tilde{w}_i,a_j,\tilde{w}_j))}_{\text{if they have a kid}}$$

where β is the discount factor, δ is the probability of loosing a job (which can happen to either of the two partners in a couple), ξ is the probability of retiring, and π^p is the probability of having a kid for a married agent. Note that $u\left((1-\tau)\frac{w(a_i)+w(a_j)}{2}\right)$ is the utility over half of the sum of the two labor income of the spouses, net of the labor income taxes given by τ.

Secondly,for a couple with one of the two agents employed we can be in one of two cases. First, if agent i is unemployed and agent j is employed, the value function is given by:

$$\beta UW(a_i, a_j, \tilde{w}_j) = \underbrace{u\left(\frac{b(a_i) + (1-\tau)w(a_j)}{2}\right)}_{\text{instantaneous utility}} + \underbrace{\delta(UU(a_i, a_j) - UW(a_i, a_j, \tilde{w}_j))}_{\text{if spouse gets unemployed}}$$
$$+ \underbrace{\lambda_u \int \max\{WW(a_i, \tilde{w}_i, a_j, \tilde{w}_j) - UW(a_i, a_j, \tilde{w}_j), 0\} dF(\tilde{w}_i)}_{\text{if she finds job}}$$
$$+ \underbrace{\xi(0 - UW(a_i, a_j, \tilde{w}_j))}_{\text{if they retire}} + \underbrace{\pi^p(UW(a_i, a_j, \tilde{w}_j; \{kid = 1\}) - UW(a_i, a_j, \tilde{w}_j))}_{\text{if they have a kid}}$$

whereas if agent i is employed and agent j is unemployed we obtain:

$$\beta WU(a_i, \tilde{w}_i, a_j) = \underbrace{u\left(\frac{(1-\tau)w(a_i) + b(a_j)}{2}\right)}_{\text{instantaneous utility}} + \underbrace{\delta(UU(a_i, a_j) - WU(a_i, \tilde{w}_i, a_j))}_{\text{if she gets unemployed}}$$
$$+ \underbrace{\lambda_u \int \max\{WW(a_i, \tilde{w}_i, a_j, \tilde{w}_j) - WU(a_i, \tilde{w}_i, a_j), 0\} dF(\tilde{w}_j)}_{\text{if spouse finds job}}$$
$$+ \underbrace{\xi(0 - WU(a_i, \tilde{w}_i, a_j))}_{\text{if they retire}} + \underbrace{\pi^p(WU(a_i, \tilde{w}_i, a_j; \{kid = 1\}) - WU(a_i, \tilde{w}_i, a_j))}_{\text{if they have a kid}}$$

In both cases, β is the discount factor, δ is the probability of loosing a job, λ_u is the probability to find a job and exit unemployment, ξ is the probability of retiring, and π^p is the probability of having a kid for a married agent. Note that $u\left(\frac{(1-\tau)w(a_i)+b(a_j)}{2}\right)$ and $u\left(\frac{b(a_i)+(1-\tau)w(a_j)}{2}\right)$ are the utility over half of the sum of the labor income of one spouse – net of the labor income taxes given by τ – and the unemployment benefit of the other spouse. As for a single individual, we have that $b(a_i) = b * e^{a_i}$ and depends on the idiosyncratic productivity of the agent a_i (similarly for an individual j characterized by a_j) and the proportional term b set by the government, which is pinned down numerical in the calibration exercise. Finally, $dF(\tilde{w}_j)$ is the distribution of job wages.

Moreover, when both agents are outLF their value function is simply given by:

$$\beta OO(a_i, a_j) = \underbrace{u\,(h)}_{\text{instantaneous utility}} + \underbrace{\xi(0 - OO(a_i, a_j))}_{\text{if they retire}} + \underbrace{\pi^p(OO(a_i, a_j; \{kid = 1\}) - OO(a_i, a_j))}_{\text{if they have a kid}}$$

where β is the discount factor, ξ is the probability of retiring, and π^p is the probability of having a kid for a married agent. Note that $u(h)$ is the utility over home production, which is the same across both individuals within the couple.

Instead, the value function of a couple with both agents unemployed is as follows:

$$
\begin{aligned}
\beta UU(a_i,a_j) = & \underbrace{u\left(\frac{b(a_i)+b(a_j)}{2}\right)}_{\text{instantaneous utility}} \\
& + \underbrace{\lambda_u \int \max\{WU(a_i,\tilde{w}_i,a_j) - UU(a_i,a_j), 0\} dF(\tilde{w}_i)}_{\text{if she finds job}} \\
& + \underbrace{\lambda_u \int \max\{UW(a_i,a_j,\tilde{w}_j) - UU(a_i,a_j), 0\} dF(\tilde{w}_j)}_{\text{if spouse finds job}} \\
& + \underbrace{\xi(0 - UU(a_i,a_j))}_{\text{if they retire}} + \underbrace{\pi^p(UU(a_i,a_j;\{kid=1\}) - UU(a_i,a_j))}_{\text{if they have a kid}}
\end{aligned}
$$

where β is the discount factor, λ_u is the probability to find a job and exit unemployment for either of the spouses, ξ is the probability of retiring, and π^p is the probability of having a kid for a married agent. Note that $u\left(\frac{b(a_i)+b(a_j)}{2}\right)$ is the utility over half of the sum of the unemployment benefits of the two agents within the couple. As for a single individual, we have that $b(a_i) = b * e^{a_i}$ and depends on the idiosyncratic productivity of the agent a_i (similarly for an individual j characterized by a_j) and the proportional term b set by the government, which is pinned down numerical in the calibration exercise. Finally, $dF(\tilde{w}_i)$ and $dF(\tilde{w}_j)$ are the distributions of job wages.

Parallel to that, the value function of a couple with one of the two agents outLF and the other one working can be one of the two following alternatives:

$$
\begin{aligned}
\beta OW(a_i,a_j,\tilde{w}_j) = & \underbrace{u\left(\frac{h+(1-\tau)w(a_j)}{2}\right)}_{\text{instantaneous utility}} + \underbrace{\delta(OU(a_i,a_j) - OW(a_i,a_j,\tilde{w}_j))}_{\text{if spouse gets unemployed}} \\
& + \underbrace{\xi(0 - OW(a_i,a_j,\tilde{w}_j))}_{\text{if they retire}} + \underbrace{\pi^p(OW(a_i,a_j,\tilde{w}_j;\{kid=1\}) - OW(a_i,a_j,\tilde{w}_j))}_{\text{if they have a kid}}
\end{aligned}
$$

if agent i is outLF and agent j is employed. Or:

$$
\begin{aligned}
\beta WO = (a_i,\tilde{w}_i,a_j) = & \underbrace{u\left(\frac{(1-\tau)w(a_i)+h}{2}\right)}_{\text{instantaneous utility}} + \underbrace{\delta(UO(a_i,a_j) - WO(a_i,\tilde{w}_i,a_j))}_{\text{if spouse gets unemployed}} \\
& + \underbrace{\xi(0 - WO(a_i,\tilde{w}_i,a_j))}_{\text{if they retire}} + \underbrace{\pi^p(WO(a_i,\tilde{w}_i,a_j;\{kid=1\}) - WO(a_i,\tilde{w}_i,a_j)))}_{\text{if they have a kid}}
\end{aligned}
$$

if agent j is outLF and agent i is employed. In both cases, β is the discount factor, δ is the probability to loose the job, ξ is the probability of retiring, and π^p is the probability of having a kid for a married agent. Note that $u\left(\frac{h+(1-\tau)w(a_j)}{2}\right)$ and $u\left(\frac{(1-\tau)w(a_i)+h}{2}\right)$ are the utility over half of the sum of the labor income of one spouse – net of the labor income taxes given by τ – and the home production of the spouse outLF. Finally, we have the two last cases, which refer to couples where one agent is unemployed and the other one is out of the labor force. First, we have:

$$\beta OU(a_i,a_j) = \underbrace{u\left(\frac{h+b(a_j)}{2}\right)}_{\text{instantaneous utility}} + \underbrace{\lambda_u \int \max\{OW(a_i,a_j,\tilde{w}_j) - OU(a_i,a_j),0\}dF(\tilde{w}_j)}_{\text{if spouse finds job}}$$
$$+ \underbrace{\xi(0 - OU(a_i,a_j))}_{\text{if they retire}} + \underbrace{\pi^p(OU(a_i,a_j;\{kid=1\}) - OU(a_i,a_j))}_{\text{if they have a kid}}$$

which is the value function if agent i is outLF and agent j is unemployed. Then we have:

$$\beta UO(a_i,a_j) = \underbrace{u\left(\frac{b(a_i)+h}{2}\right)}_{\text{instantaneous utility}} + \underbrace{\lambda_u \int \max\{WO(a_i,\tilde{w}_i,a_j) - UO(a_i,a_j),0\}dF(\tilde{w}_i)}_{\text{if spouse finds job}}$$
$$+ \underbrace{\xi(0 - UO(a_i,a_j))}_{\text{if they retire}} + \underbrace{\pi^p(UO(a_i,a_j;\{kid=1\}) - UO(a_i,a_j))}_{\text{if they have a kid}}$$

which is the value function if agent j is outLF and agent i is unemployed. In both cases, β is the discount factor, λ_u is the probability to find a job and exit unemployment, ξ is the probability of retiring, and π^p is the probability of having a kid for a married agent. Note that $u\left(\frac{h+b(a_j)}{2}\right)$ and $u\left(\frac{b(a_i)+h}{2}\right)$ are the utilities over half of the sum of the unemployment benefit and the home production of the partners within this type of couple. As for a single individual, we have that $b(a_i) = b * e^{a_i}$ and depends on the idiosyncratic productivity of the agent a_i (similarly for an individual j characterized by a_j) and the proportional term b set by the government, which is pinned down numerical in the calibration exercise. Finally, $dF(\tilde{w}_j)$ and $dF(\tilde{w}_i)$ are the distributions of job wages.

Bibliography

Abbott, B., Gallipoli, G., Meghir, C., and Violante, G. L. (2019). Education policy and intergenerational transfers in equilibrium. *Journal of Political Economy*, 127(6):2569–2624.

Alati, A. (2020). *Essays on firms heterogeneity and business cycles*. PhD thesis, The London School of Economics and Political Science (LSE).

Albanesi, S. and Nosal, J. (2018). Insolvency after the 2005 bankruptcy reform. Technical report, National Bureau of Economic Research.

Alon, T., Bachas, N., and Wong, A. (2021). Debt, Human Capital Accumulation, and the Allocation of Talent. Working paper.

Ambrose, B. W., Cordell, L., and Ma, S. (2015). The impact of student loan debt on small business formation. *Available at SSRN 2417676*.

Antonovics, K. and Town, R. (2004). Are all the good men married? uncovering the sources of the marital wage premium. *American Economic Review*, 47(2):317–321.

Atkeson, A. and Burstein, A. (2008). Pricing-to-market, trade costs, and international relative prices. *American Economic Review*, 98(5):1998–2031.

Auclert, A. (2019). Monetary policy and the redistribution channel. *American Economic Review*, 109(6):2333–67.

Avery, C. and Turner, S. (2012). Student loans: Do college students borrow too much–or not enough? *Journal of Economic Perspectives*, 26(1):165–92.

Baley, I. and Blanco, A. (2019). Firm uncertainty cycles and the propagation of nominal shocks. *American Economic Journal: Macroeconomics*, 11(1):276–337.

Baqaee, D., Farhi, E., and Sangani, K. (2021). The supply-side effects of monetary policy. Technical report, National Bureau of Economic Research.

Batty, M., Bricker, J., Briggs, J., Holmquist, E., McIntosh, S., Moore, K., Nielsen, E., Reber, S., Shatto, Molly, S. K., Sweeney, T., and Henriques Volz, A. (2019). Introducing the Distributional Financial Accounts of the United States. *Finance and Economics Discussion Series 2019-017. Washington: Board of Governors of the Federal Reserve System*.

Berkner, L. K. (2000). *Trends in undergraduate borrowing: Federal student loans in 1989-90, 1992-93, and 1995-96*. DIANE Publishing.

Bilbiie, F. O., Ghironi, F., and Melitz, M. J. (2012). Endogenous entry, product variety, and business cycles. *Journal of Political Economy*, 120(2):304–345.

Bilbiie, F. O., Ghironi, F., Melitz, M. J., Midrigan, V., and Rotemberg, J. J. (2007). Monetary policy and business cycles with endogenous entry and product variety [with comments and discussion]. *NBER Macroeconomics Annual*, 22:299–379.

Bils, M., Klenow, P. J., and Malin, B. A. (2018). Resurrecting the role of the product market wedge in recessions. *American Economic Review*, 108(4-5):1118–46.

Blandin, A. and Herrington, C. (2020). Family heterogeneity, human capital investment, and college attainment. *Human Capital Investment, and College Attainment (September 14, 2020)*.

Borella, M., DeNardi, M., and Yang, F. (2019). Are marriage-related taxes and social security benefits holding back female labor supply? *NBER working paper no. 26097*.

Brown, M., Haughwout, A., Lee, D., Scally, J., and Van Der Klaauw, W. (2015). Measuring student debt and its performance. *Student loans and the dynamics of debt*, pages 37–52.

Buera, F. J., Kaboski, J. P., and Shin, Y. (2011). Finance and Development: A Tale of Two Sectors. *American Economic Review*, 101(5):1964–2002.

Buera, F. J., Kaboski, J. P., and Shin, Y. (2021a). The macroeconomics of microfinance. *The Review of Economic Studies*, 88(1):126–161.

Buera, F. J., Kaboski, J. P., and Townsend, R. M. (2021b). From micro to macro development.

Buera, F. J. and Shin, Y. (2013). Financial Frictions and the Persistence of History: A Quantitative Exploration. *Journal of Political Economy*, 121(2):221–272.

Burstein, A., Carvalho, V. M., and Grassi, B. (2020). Bottom-up markup fluctuations. Technical report, National Bureau of Economic Research.

Cagetti, M. and De Nardi, M. (2006). Entrepreneurship, frictions, and wealth. *Journal of Political Economy*, 114(5):835–870.

Cai, Z. and Heathcote, J. (2022). College tuition and income inequality. *American Economic Review*, 112(1):81–121.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2015). Optimal data-driven regression discontinuity plots. *Journal of the American Statistical Association*, 110(512):1753–1769.

Calvo, P. A., Lindenlaub, I., and Reynoso, A. (2021). Marriage market and labor market sorting. Technical report, National Bureau of Economic Research.

Catherine, S. and Yannelis, C. (2020). The distributional effects of student loan forgiveness. Technical report, National Bureau of Economic Research.

Chatterjee, S. and Ionescu, F. (2012). Insuring student loans against the financial risk of failing to complete college. *Quantitative Economics*, 3(3):393–420.

Chetty, R., Hendren, N., Kline, P., and Saez, E. (2014). Where is the land of opportunity? the geography of intergenerational mobility in the united states. *The Quarterly Journal of Economics*, 129(4):1553–1623.

Choi, S. and Valladares-Esteban, A. (2018). The marriage unemployment gap. *The B.E. Journal of Macroeconomics*, 18(1).

Choi, S. and Valladares-Esteban, A. (2019). On households and unemployment insurance. *Quantitative Economics*, 11:437–469.

Clementi, G. L. and Palazzo, B. (2016). Entry, Exit, Firm Dynamics, and Aggregate Fluctuations. *American Economic Journal: Macroeconomics*, 8(3):1–41.

Cloyne, J., Ferreira, C., Froemel, M., and Surico, P. (2018). Monetary policy, corporate finance and investment. Technical report, National Bureau of Economic Research.

Cohen-Cole, E., Duygan-Bump, B., and Montoriol-Garriga, J. (2013). Who gets credit after bankruptcy and why? an information channel. *Journal of Banking & Finance*, 37(12):5101–5117.

Colas, M., Findeisen, S., and Sachs, D. (2021). Optimal need-based financial aid. *Journal of Political Economy*, 129(2):492–533.

Daruich, D. (2018). The macroeconomic consequences of early childhood development policies. *Available at SSRN 3265081*.

Daruich, D. and Kozlowski, J. (2020). Explaining intergenerational mobility: The role of fertility and family transfers. *Review of Economic Dynamics*, 36:220–245.

Davis, S. J., Haltiwanger, J., Jarmin, R., Miranda, J., Foote, C., and Nagypal, E. (2006). Volatility and dispersion in business growth rates: Publicly traded versus privately held firms [with comments and discussion]. *NBER macroeconomics annual*, 21:107–179.

De Loecker, J., Eeckhout, J., and Unger, G. (2020). The rise of market power and the macroeconomic implications. *The Quarterly Journal of Economics*, 135(2):561–644.

De Loecker, J. and Warzynski, F. (2012). Markups and firm-level export status. *American economic review*, 102(6):2437–71.

De Nardi, M., Fella, G., and Paz-Pardo, G. (2020). Nonlinear household earnings dynamics, self-insurance, and welfare. *Journal of the European Economic Association*, 18(2):890–926.

Decker, R., Haltiwanger, J., Jarmin, R., and Miranda, J. (2014). The role of entrepreneurship in us job creation and economic dynamism. *Journal of Economic Perspectives*, 28(3):3–24.

Di Maggio, M., Kalda, A., and Yao, V. (2019). Second chance: Life without student debt. Technical report, National Bureau of Economic Research.

Dinlersoz, E., Kalemli-Ozcan, S., Hyatt, H., and Penciakova, V. (2018). Leverage over the life cycle and implications for firm growth and shock responsiveness. Technical report, National Bureau of Economic Research.

Doepke, M. and Gaetani, R. (2020). Why didn't the college premium rise everywhere? employment protection and on-the-job investment in skills. Technical report, National Bureau of Economic Research.

Doepke, M. and Tertilt, M. (2016). Families in Macroeconomics. In Taylor, J. B. and Uhlig, H., editors, *Handbook of Macroeconomics*, chapter 23, pages 1789–1891. Elsevier.

Dunne, T., Roberts, M. J., and Samuelson, L. (1989). The growth and failure of us manufacturing plants. *The Quarterly Journal of Economics*, 104(4):671–698.

Edmond, C., Midrigan, V., and Xu, D. Y. (2018). How costly are markups? Technical report, National Bureau of Economic Research.

Ek, S. and Holmlund, B. (2010). Family job search, wage bargaining, and optimal unemployment insurance. *The B.E. Journal of Economic Analysis and Policy*, 1(47).

Fabiani, A., Falasconi, L., and Heineken, J. (2020). Monetary policy and corporate debt maturity.

Fairlie, R. W. and Robb, A. M. (2009). Gender differences in business performance: evidence from the characteristics of business owners survey. *Small Business Economics*, 33(4):375–395.

Flabbi, L. and Finn, C. (2015). Simultaneous search in the labor and marriage markets with endogenous schooling decisions. *BGSE Working Paper*.

Flabbi, L. and Mabli, J. (2018). Household search or individual search: Does it matter? *Journal of Labor Economics*, 36(1):1–46.

Folch, M. and Mazzone, L. (2020). Go big or buy a home.

Fortin, N. M. (2006). Higher-education policies and the college wage premium: Cross-state evidence from the 1990s. *American Economic Review*, 96(4):959–987.

Galí, J. (2015). *Monetary policy, inflation, and the business cycle: an introduction to the new Keynesian framework and its applications*. Princeton University Press.

Gali, J., Gertler, M., and Lopez-Salido, J. D. (2007). Markups, gaps, and the welfare costs of business fluctuations. *The review of economics and statistics*, 89(1):44–59.

Gemici, A. (2011). Family migration and labor market outcomes. *Working Paper*.

Gertler, M. and Gilchrist, S. (1994). Monetary policy, business cycles, and the behavior of small manufacturing firms. *The Quarterly Journal of Economics*, 109(2):309–340.

Gicheva, D. and Thompson, J. (2015). The effects of student loans on long-term household financial stability. *Student loans and the dynamics of debt*, pages 287–316.

Gilchrist, S., Schoenle, R., Sim, J., and Zakrajšek, E. (2017). Inflation dynamics during the financial crisis. *American Economic Review*, 107(3):785–823.

Goldin, C. and Katz, L. F. (2010). *The race between education and technology*. harvard university press.

Golosov, M. and Lucas, R. E. (2007). Menu costs and phillips curves. *Journal of Political Economy*, 115(2):171–199.

Gordon, G. and Hedlund, A. (2020). Accounting for tuition increases across us colleges. *Available at SSRN 3720860*.

Greenwood, J., Guner, N., Kocharkov, G., and Santos, C. (2014). Marry your like: Assortative mating and income inequality. *The American Economic Review*, 104.

Guler, B., Guvenen, F., and Violante, G. (2012). Joint-search theory: New opportunities and new frictions. *Journal of Monetary Economics*, 59(4):352–369.

Guner, N., Kaygusuz, R., and Ventura, G. (2012). Taxation and household labor supply. *The Review of Economic Studies*, 79(3):1113–1149.

Guner, N., Kaygusuz, R., and Ventura, G. (2020). Child-related transfers, household labour supply, and welfare. *The Review of Economic Studies*, 87(5):2290–2321.

Guo, X., Chen, W., and Yu, A. (2016). Is college education worth it? evidence from its impacts on entrepreneurship in the united states. *Journal of Small Business & Entrepreneurship*, 28(1):1–26.

Gürkaynak, R. S., Sack, B., and Swanson, E. (2005). The sensitivity of long-term interest rates to economic news: Evidence and implications for macroeconomic models. *American economic review*, 95(1):425–436.

Guvenen, F., Kuruscu, B., Tanaka, S., and Wiczer, D. (2020). Multidimensional skill mismatch. *American Economic Journal: Macroeconomics*, 12(1):210–44.

Hall, R. E. (1988). The relation between price and marginal cost in us industry. *Journal of political Economy*, 96(5):921–947.

Haltiwanger, J., Jarmin, R. S., and Miranda, J. (2013). Who creates jobs? small versus large versus young. *Review of Economics and Statistics*, 95(2):347–361.

Heathcote, J., Storesletten, K., and Violante, G. L. (2010). The macroeconomic implications of rising wage inequality in the united states. *Journal of political economy*, 118(4):681–722.

Herkenhoff, K., Phillips, G. M., and Cohen-Cole, E. (2021). The impact of consumer credit access on self-employment and entrepreneurship. *Journal of financial economics*, 141(1):345–371.

Hershbein, B. and Hollenbeck, K. M. (2015). *Student loans and the dynamics of debt*. WE Upjohn Institute.

Hill, M. S. (1979). The wage effects of marital status and children. *The Journal of Human Resources*, 14(4):579–594.

Hong, S. (2017). Customer capital, markup cyclicality, and amplification. *FRB St. Louis Working Paper*, (2017-33).

Hsieh, C.-T. and Klenow, P. (2009). Misallocation and Manufacturing TFP in China and India. *The Quarterly Journal of Economics*, 124(4):1403–1448.

Hua, Y. (2021). The long-run effects of federal student loans on fertility and social mobility.

Huckfeldt, C. (2016). Understanding the scarring effect of recessions. *Report, Economics Department*.

Hurst, E. and Lusardi, A. (2004). Liquidity constraints, household wealth, and entrepreneurship. *Journal of political Economy*, 112(2):319–347.

Ionescu, F. (2009). The federal student loan program: Quantitative implications for college enrollment and default rates. *Review of Economic dynamics*, 12(1):205–231.

Ionescu, F. (2011). Risky human capital and alternative bankruptcy regimes for student loans. *Journal of Human Capital*, 5(2):153–206.

Ionescu, F. and Simpson, N. (2016). Default risk and private student loans: Implications for higher education policies. *Journal of Economic Dynamics and Control*, 64:119–147.

Iuliano, J. (2012). A empirical assessment of student loan discharges and the undue hardship standard. *Am. Bankr. LJ*, 86:495.

Jarociński, M. and Karadi, P. (2020). Deconstructing monetary policy surprises—the role of information shocks. *American Economic Journal: Macroeconomics*, 12(2):1–43.

Jarosch, G. (2021). Searching for job security and the consequences of job loss. Technical report, National Bureau of Economic Research.

Jeenas, P. (2019). Firm balance sheet liquidity, monetary policy shocks, and investment dynamics.

Ji, Y. (2021). Job search under debt: Aggregate implications of student loans. *Journal of Monetary Economics*, 117:741–759.

Jiang, H. and Sohail, F. (2017). Skill biased entrepreneurial decline. Technical report, Working paper, September 2017. 5.

Jones, J. B. and Yang, F. (2016). Skill-biased technical change and the cost of higher education. *Journal of Labor Economics*, 34(3):621–662.

Jordà, Ò. (2005). Estimation and inference of impulse responses by local projections. *American economic review*, 95(1):161–182.

Juhn, C. and McCue, K. (2016). Evolution of the marriage earnings gap for women. *American Economic Review*, 106(5):252–256.

Kaboski, J. P. and Townsend, R. M. (2011). A structural evaluation of a large-scale quasi-experimental microfinance initiative. *Econometrica*, 79(5):1357–1406.

Kaplan, G., Moll, B., and Violante, G. L. (2018). Monetary policy according to hank. *American Economic Review*, 108(3):697–743.

Kaplan, S. N. and Rauh, J. (2013). It's the market: The broad-based rise in the return to top talent. *Journal of Economic Perspectives*, 27(3):35–56.

Kerdelhué, L. (2021). Scarred young entrepreneurs: the effects of education policies.

Kim, H. and Kim, J. H. (2022). Student debt, college tuition, and wage inequality.

Kiyotaki, N. and Moore, J. (1997). Credit cycles. *Journal of political economy*, 105(2):211–248.

Klenow, P. J. and Willis, J. L. (2016). Real rigidities and nominal price changes. *Economica*, 83(331):443–472.

Korenman, S. and Neumark, D. (1991). Does marriage really make men more productive? *Journal of Human Resources*, 26(2):282–307.

Kozeniauskas, N. (2018). What's driving the decline in entrepreneurship. *Unpublished paper. New York University, New York, NY*.

Krishnan, K. and Wang, P. (2019). The cost of financing education: can student debt hinder entrepreneurship? *Management Science*, 65(10):4522–4554.

Lagakos, D., Moll, B., Porzio, T., Qian, N., and Schoellman, T. (2018). Life cycle wage growth across countries. *Journal of Political Economy*, 126(2):797–849.

Lee, Y. and Mukoyama, T. (2015). Productivity and Employment Dynamics of US Manufacturing Plants. *Economics Letters*, 136:190–193.

Lerner, J. and Malmendier, U. (2013). With a little help from my (random) friends: Success and failure in post-business school entrepreneurship. *The Review of Financial Studies*, 26(10):2411–2452.

Levy, D., Bergen, M., Dutta, S., and Venable, R. (1997). The magnitude of menu costs: direct evidence from large us supermarket chains. *The Quarterly Journal of Economics*, 112(3):791–824.

Lochner, L. and Monge-Naranjo, A. (2016). Student loans and repayment: Theory, evidence, and policy. In *Handbook of the Economics of Education*, volume 5, pages 397–478. Elsevier.

Looney, A. and Yannelis, C. (2015a). A crisis in student loans?: How changes in the characteristics of borrowers and in the institutions they attended contributed to rising loan defaults. *Brookings Papers on Economic Activity*, 2015(2):1–89.

Looney, A. and Yannelis, C. (2015b). Is high student loan debt always a problem. *SIEPR Policy Brief*.

Lucas, R. E. (1978). On the size distribution of business firms. *The Bell Journal of Economics*, pages 508–523.

Luo, M. and Mongey, S. (2019). Assets and job choice: Student debt, wages and amenities. Technical report, National Bureau of Economic Research.

Mankart, J. and Oikonomou, R. (2017). Household search and the aggregate labour market. *The Review of Economic Studies*, 84(4):1735–1788.

Matsuda, K. (2020). Optimal timing of college subsidies: Enrollment, graduation, and the skill premium. *European Economic Review*, 129:103549.

Matsuda, K. and Mazur, K. (2022). College education and income contingent loans in equilibrium. *Journal of Monetary Economics*.

McCall, J. (1970). Economics of Information and Job Search. *The Quarterly Journal of Economics*, 84(1):113–126.

McConnell, B. and Valladares-Esteban, A. (2020). On the marriage wage premium. *Working Paper*.

McKay, A., Nakamura, E., and Steinsson, J. (2016). The power of forward guidance revisited. *American Economic Review*, 106(10):3133–58.

Meier, M. and Reinelt, T. (2020). Monetary policy, markup dispersion, and aggregate tfp.

Mezza, A., Ringo, D., Sherlund, S., and Sommer, K. (2020). Student loans and homeownership. *Journal of Labor Economics*, 38(1):215–260.

Mezza, A., Ringo, D., and Sommer, K. (2021). Student loans, access to credit and consumer financial behavior.

Michelacci, C. and Schivardi, F. (2020). Are they all like bill, mark, and steve? the education premium for entrepreneurs. *Labour Economics*, 67:101933.

Midrigan, V. and Xu, D. Y. (2014). Finance and Misallocation: Evidence from Plant-Level Data. *American Economic Review*, 104(2):422–458.

Mongey, S. (2017). Market structure and monetary non-neutrality. *Job market paper. http://www. simonmongey. com/uploads/6/5/6/6/65665741/mongey_market_structure_monetary_nonneutrality_draft.pdf*.

Morazzoni, M. and Sy, A. (2021). Female entrepreneurship, financial frictions and capital misallocation in the us.

Mueller, H. M. and Yannelis, C. (2019). The rise in student loan defaults. *Journal of Financial Economics*, 131(1):1–19.

Nagypal, E. (2007). Labor-market fluctuations and on-the-job search. *Manuscript, Northwestern University*.

Nakosteen, R. A. and Zimmer, M. A. (1987). Marital status and earnings of young men: A model with endogenous selection. *The Journal of Human Resources*, 22(2):248–268.

Neal, D. (1995). Industry-specific human capital: Evidence from displaced workers. *Journal of labor Economics*, 13(4):653–677.

Nekarda, C. J. and Ramey, V. A. (2020). The cyclical behavior of the price-cost markup. *Journal of Money, Credit and Banking*, 52(S2):319–353.

Ortigueira, S. and Siassi, N. (2013). How Important is Intrahousehold Risk Sharing for Savings and Labor Supply? *Journal of Monetary Economics*, 60(6).

Ortigueira, S. and Siassi, N. (2020). The u.s. tax-transfer system and low-income households: Savings, labor supply, and household formation. *Mimeo*.

Ottonello, P. and Winberry, T. (2020). Financial heterogeneity and the investment channel of monetary policy. *Econometrica*, 88(6):2473–2502.

Pilossoph, L. and Wee, S. L. (2020). Assortative matching and household income inequality: A structural approach. *Working Paper*.

Pilossoph, L. and Wee, S. L. (2021). Household search and the marital wage premium. *American Economic Journal: Macroeconomics*, 13(4):55–109.

Poschke, M. (2013). Who becomes an entrepreneur? labor market prospects and occupational choice. *Journal of Economic Dynamics and Control*, 37(3):693–710.

Pugsley, B. W., Sedlacek, P., and Sterk, V. (2019). The nature of firm growth. *Available at SSRN 3086640*.

Quadrini, V. (2009). Entrepreneurship in macroeconomics. *Annals of Finance*, 5(3):295–311.

Queiró, F. (2021). Entrepreneurial human capital and firm dynamics. *Available at SSRN 3280925*.

Robb, A. M. and Robinson, D. T. (2014). The capital structure decisions of new firms. *The Review of Financial Studies*, 27(1):153–179.

Salgado, S. (2020). Technical change and entrepreneurship. *Available at SSRN 3616568*.

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

Shimer, R. (2005). The cyclical behavior of equilibrium unemployment and vacancies. *American Economic Review*, 95(1):25–49.

Taylor, J. B. (1999). The robustness and efficiency of monetary policy rules as guidelines for interest rate setting by the european central bank. *Journal of Monetary Economics*, 43(3):655–679.

Van der Sluis, J., Van Praag, M., and Vijverberg, W. (2008). Education and entrepreneurship selection and performance: A review of the empirical literature. *Journal of economic surveys*, 22(5):795–841.

Vardishvili, O. (2020). The macroeconomic cost of college dropouts. *Available at SSRN 3755800*.

Wei, C. C. and Berkner, L. (2008). Trends in undergraduate borrowing ii: Federal student loans in 1995-96, 1999-2000, and 2003-04. postsecondary education descriptive analysis report. nces 2008-179. *National Center for Education Statistics*.

Wong, A. (2019). Refinancing and the transmission of monetary policy to consumption. *Unpublished manuscript*, 20.

Yannelis, C. (2016). Strategic default on student loans. *New York University, Working Paper*.

Zbaracki, M. J., Ritson, M., Levy, D., Dutta, S., and Bergen, M. (2004). Managerial and customer costs of price adjustment: direct evidence from industrial markets. *Review of Economics and statistics*, 86(2):514–533.

Zucman, G. (2019). Global Wealth Inequality. *Annual Review of Economics*, 11:109–138.