

Essays on Intangible Investment

A DISSERTATION

SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA

BY

Thomas J. May

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

Doctor of Philosophy

Ellen R. McGrattan, Advisor

July, 2024

© Thomas J. May 2024

ALL RIGHTS RESERVED

Acknowledgements

I am deeply grateful to my advisor Dr. Ellen McGrattan, and my co-authors Dr. Anmol Bhandari, Dr. Tobey Kass for their advice and feedback. I would also like to acknowledge the entire University of Minnesota Department of Economic's faculty and economists of the Federal Reserve Bank of Minneapolis. Especially, Dr. Kjetil Storesletten and Dr. Todd Schoellman. I am also thankful to my fellow University of Minnesota Economics graduate students, particularly Dr. Fil Bablievsky and Dr. Shahid Ahmed.

Dedication

To my mom, who has always been there for me every step of the way. To my wife, whom I love so much. To my baby boy, Jacobi.

Abstract

This dissertation attempts to deepen the literature's understanding of intangible investments. Specifically, it focuses on the impact of regulatory and fiscal policy on motivations for intangible investments. This dissertation consists of three chapters. The first chapter considers the impact of antitrust policy on research and development incentives. Policymakers have increasingly concluded that optimal antitrust policy requires looking beyond traditional static analyses and considering the dynamic effects of policy. Such analysis is challenging as limited studies exist concerning dynamic competition policy. This paper attempts to bridge this knowledge gap by developing a novel structural growth model that includes mergers and acquisitions (M&A) activity. This allows for a deeper understanding of the impact of M&A on economic growth as well as serving as a test bed for proposed antitrust policies. The model, estimated on U.S. data, shows that M&A is a double-edged sword. Blocking all M&A transactions causes a significant decrease in the growth rate. Unrestricted M&A causes a significant increase in market concentration. Both cases are welfare decreasing. The chapter then considers the balance of static vs. dynamic policy as well as a policy regime's impact on the macroeconomy.

Antitrust policy analysis requires an understanding the motives behind a proposed

transaction. The second chapter develops a novel methodology for classifying relationships between parties in a merger and acquisition transaction. Understanding the motives of an M&A transaction is essential in researching the economic impact of M&A. Given the high number of transactions per year, manually classifying every transaction is unfeasible. This paper proposes a novel methodology using a large-language model to determine if a transaction has possible horizontal or vertical linkages. With sufficient information, this model is highly accurate. Focusing on transactions involving U.S. firms, 38% of transactions were only horizontally linked, 38% were only vertically linked, 12% were both horizontally and vertically linked, and 10% had no linkages. This pattern is robust across time and presidential administrations. The resulting data can be used by any researcher studying M&A transactions.

Continuing with the theme of understanding motives behind intangible investments, the third chapter elucidates the nature of entrepreneurship by comparing life-cycle income profiles and outcomes of individuals who share similar characteristics but differ in their choice of self- or paid-employment. Results are based on U.S. administrative data from the Internal Revenue Service and Social Security Administration over the period 2000–2015 for subgroups of the population differing by gender, marital status, education, occupation, industry, cohort, and employment status. Contrary to top-coded survey evidence based on relatively small samples and short panels,

we find that entrepreneurs with at least twelve years in self-employment during our sample have significantly higher average income and steeper, more persistent, income growth profiles than their paid-employed peers with similar characteristics. Contrary to survey evidence, we find that new entrants into self-employment have higher labor incomes and lower asset incomes prior to entry relative to similar peers that do not enter. A theory of entrepreneurial choice is developed and compared to the subsample of young entrepreneurs in our data. We find that including firm-specific investment and selection under incomplete information is necessary if the theory is to match the observed income growth profiles and switching behavior for these young entrepreneurs.

Contents

Acknowledgements	i
Dedication	ii
Abstract	iii
List of Tables	xi
List of Figures	xiii
1 Innovation and Competition Policy	1
1.1 Introduction	1
1.1.1 Literature Review	6
1.1.2 Institutional Details: U.S. Antitrust Policy	9
1.2 Model	12
1.2.1 Environment	12

1.2.2	Equilibrium	26
1.2.3	Equilibrium Results	27
1.2.4	Numerical Solutions	30
1.3	Estimation	31
1.3.1	The Antitrust Policy Function	31
1.3.2	Functional Forms	35
1.3.3	Parameter Estimation	35
1.4	The Effect of M&A on Growth	37
1.4.1	Impact of M&A Across the Firm Distribution	37
1.4.2	M&A and Aggregate Growth	39
1.5	Competition Policy	40
1.5.1	The Dual Edge Sword Effect of M&A	40
1.5.2	Simple Antitrust Rules	41
1.5.3	Banning Killer Acquisitions	44
1.5.4	The Importance of Endogenous M&A for Policy Analysis	45
1.6	Conclusion	46
1.7	Tables	48
1.8	Figures	54
2	Employing Natural Language Processing To Understand M&A Mo-	
	tives	57

2.1	Introduction	57
2.2	Data on U.S. M&A Transactions	61
2.3	Employing LLMs to Study M&A Motivations	64
2.3.1	Example Transactions	66
2.3.2	Evaluating the Model’s Accuracy	70
2.4	Summary of Classification Results	72
2.5	Conclusion	74
2.6	Tables	76
2.7	Figures	81
3	On the Nature of Entrepreneurship	84
3.1	Introduction	85
3.2	Data	92
3.2.1	Sample	92
3.2.2	Income Measures	94
3.2.3	Imputations for Skills and Education	98
3.3	Measuring Returns to Employment	103
3.3.1	Some Challenges	103
3.3.2	Econometric Framework	104
3.3.3	Groups	110
3.4	Results	115

3.4.1	Entrepreneurial Incomes	116
3.4.2	Entrepreneurial Choice	132
3.4.3	Young Entrepreneurs	137
3.4.4	Comparison to existing literature	140
3.5	Implications for Theory	146
3.5.1	Model	147
3.5.2	Quantitative results	151
3.6	Conclusions	158
3.7	Tables	160
3.8	Figures	167
References		189
Appendix A. Innovation and Competition Policy Derivations		196
Appendix B. Numerical Solution Method for the Innovation and Competition Policy Model		205
B.1	High-Level Overview	205
B.2	Model Solution Steps	206
B.2.1	Construct the State Set	206
B.2.2	Solve the Cournot Game	207
B.2.3	Construct Division Linear Maps	209

B.2.4	Construct Market Linear Maps	215
B.2.5	Construct Entry Map	217
B.2.6	Computing Antitrust Blocking Probabilities	218
B.2.7	Solve for the Stationary Distribution	218
B.3	Supporting Proofs	224
Appendix C. Antitrust Policy Regression		227
C.1	Data Preparation	227
C.2	Regression Results	228
C.3	Tables	230
C.4	Figures	231

List of Tables

1.1	Outcome of Second Reviews by HHI and Δ HHI	48
1.2	Estimated Parameters and Target Moments	49
1.3	Entry Effort By Firm Productivity	50
1.4	Sources of Aggregate Growth (Percent)	51
1.5	Alternative Antitrust Policies	52
1.6	Simple Antitrust Rules	53
2.1	M&A Transaction Summary Statistics	76
2.2	Classification Confusion Matrix	77
2.3	Classification Precision and Recall	78
2.4	Count Shares of M&A Transaction	79
2.5	Value Shares of M&A Transaction	80
3.1	Main Sample Summary Statistics	160
3.1	Main Sample Summary Statistics (Continued)	161
3.2	Largest Contributors to Attached Employee Growth Gap	162

3.3	Transition Probabilities for Incomes in Levels	163
3.4	Transition Probabilities for Residual Growth Rates	164
3.5	Income Shares Held by Each Employment Group	165
3.6	Model Parameters	166
C.1	$\hat{\mathbb{P}}(B S, \text{hhi}, \Delta\text{hhi})$ Regression Results	230

List of Figures

1.1	Venture Capital Exits By Type of Exit	54
1.2	Frequency of Second Antitrust Reviews and Challenges	55
1.3	Inverse Cumulative Distribution for Change in Improvement Effort	56
2.1	M&A Transaction Counts and Median Value By Year	81
2.2	Shares of Transactions	82
2.3	Transaction Shares by Presidential Administration	83
3.1	Estimated Time Effects Relative to Average Income	167
3.2	Income and Growth Profiles	168
3.3	Growth Profiles by Gender	169
3.4	Growth Profiles by Marital Status	170
3.5	Growth Profiles by Education	171
3.6	Growth Profiles by Cognitive Skill	172
3.7	Growth Profiles by Interpersonal Skill	173
3.8	Growth Profiles by Manual Skill	174

3.9	Growth Profiles by Industry	175
3.10	Growth Profiles of Mostly Switchers and Any Non-Employed	176
3.11	Age-Over-Age Growth in Incomes	177
3.12	Age-Over-Age Growth in Regression Residuals	178
3.13	Self-Employment Switching Rates By Age	179
3.14	Self-Employment Switching Rates By Year	180
3.15	Differences in Past Incomes Between Current and Future Switchers .	181
3.15	Differences in Past Incomes Between Current and Future Switchers (Cont.)	182
3.16	Differences in Incomes Between Switchers and Non-switchers	183
3.17	Growth Differentials for Young Entrepreneurs	184
3.18	Empirical Moments, IRS versus CPS	185
3.19	Growth Profiles for Top and Bottom 25% Income Ranks	186
3.20	Growth Differentials for Young Entrepreneurs	187
3.21	Model Predictions as Intangible Revenue Varied	188
C.1	Estimated $\hat{\mathbb{P}}(B S, \text{hhi}, \Delta\text{hhi})$	231

Chapter 1

Innovation and Competition Policy

1.1 Introduction

In 2004, a company named MapQuest dominated the world of online mapping and navigation. At the same time, a small tech company called Where 2 Technologies had also developed a competitor product called Expedition — one that would prove vastly superior to the existing products. Where 2 secured an initial funding deal from a venture capital firm, but any hope of a deal was lost as Yahoo! entered the market the next day. The consensus from venture firms was that it would require too much investment and risk to enable Where 2 to compete with an internet giant like Yahoo. However, one of those VC funds connected Where 2 with another client interested in bringing Where 2's product to the market and out-compete Yahoo. That client was

Google. Expedition was re-branded as Google Maps. It is fair to say that, without being acquired, the return on investment for Where 2 Technologies' product would have been scarce.

This story is not just an anecdote of one tech startup. Mergers and Acquisitions (M&A) have become the primary channel for startup companies to receive a return on their investments. Figure 1.1 displays methods of VC firm exits, the point where venture capital firms exit their investments and receive returns on their investment. It shows that from 2010 to 2019, on average, 78% of VC firms exited via acquisitions compared to 8%, which were exits through public offerings. This trend is a reversal of the VC exit outcomes of the 1990s. From the acquirers' perspective, M&A is a significant channel of investment. Data from SDC platinum, the most comprehensive dataset on U.S. M&A transactions, shows that M&A transaction values, on average, exceed \$300 billion a year. However, M&A is a channel rarely considered in the endogenous growth literature. By excluding M&A activity, the literature is missing a significant incentive for investment and a central mechanism for firm growth.

In addition to being a source of significant investment incentives, M&A is also one of the primary methods for firms to consolidate markets and gain monopoly power. Monopolies and antitrust issues have always been a traditional concern of the government. However, recent developments in market consolidation and the rise of large

tech companies have brought increased scrutiny. In 2020, the House Judiciary Committee’s antitrust panel issued a congressional report following an investigation into the big four tech companies: Amazon, Apple, Google, and Facebook. In particular, the report raised concerns that the tech companies were no longer confined to their original scope and instead had become multi-headed corporations with arms in many different markets. Another key concern was killer acquisitions, as documented by [Cunningham et al. \[2021\]](#), where a company attempts to maintain its current level of market power by buying a competitor within the same market with the sole purpose of retiring the competing product and shutting down its further development.

From the perspective of an antitrust regulator, fighting markup-induced misallocation while minimizing the distortion on R&D incentives poses a difficult policy problem. As [Crandall and Winston \[2003\]](#) points out, ”the substantial and growing challenges of formulating and implementing effective antitrust policies in a new economy characterized by dynamic competition, rapid technological change and important intellectual property.” Given the existing literature, these challenges are difficult to address because the study of antitrust dynamic competition policy is incredibly limited. This sentiment is echoed in the remarks of J. Thomas Rosch, a former commissioner of the Federal Trade Commission:

If you were to get together a group of antitrust and patent experts, everyone would likely agree broadly that antitrust and intellectual property are complementary in that both areas of law seek to protect and encourage innovation. When it comes to antitrust law, however, promoting innovation is good in theory, but hard in practice.

Thus, the goal of this paper is to help develop a useful tool for the consideration of different dynamic antitrust concerns and policies.

To accomplish that goal, this paper develops a structural model of R&D and M&A to accomplish this task. Within the model, firms can be multi-product. They use R&D investments to both improve the products currently in their portfolio as well as create new products in order to enter into new markets. The M&A block allows the firms to buy the product lines of other firms. This captures M&A's incentive for R&D and the ability for M&A to increase a firm's market power. M&A in the model is solely horizontal in nature. While other forms of M&A are incredibly interesting, antitrust enforcement is primarily focused on horizontal mergers. The Federal Trade Commission's *Horizontal Merger Investigation Data, Fiscal Years 1996-2011* reveals 264 investigation's theory of potential violations were horizontal and 28 were vertical. So this chapter chooses to focus on the horizontal channel.

After estimating the model on U.S. economic data, this chapter shows the overall impact of M&A on growth is positive. In an alternative economy where there was no

M&A, the model estimates the economic growth rate would be 0.35pp lower than the estimated baseline growth rate. This decrease comes from two sources: the inability to combine existing products to create new products and a decreased incentive to create new products to enter into new markets. The decrease in entry rates is most prominent in low-productivity firms. They experience a decrease in entry rates by 53%.

While this shows the importance of M&A in investment decisions, M&A is a double-edged sword. Unrestricted M&A is also welfare decreasing compared to the estimated baseline model. While it increases growth, it also causes a massive increase in concentration and misallocation. Thus, the antitrust policymaker must balance its effect on growth with its effect on market concentration.

This chapter then considers the impact of antitrust policy through simple rules. This considers the impact of a proposed transaction on the transaction's relevant market. The model shows that the policy maker who focuses on dynamic outcomes, the long-run effects of the M&A transaction, is significantly welfare improving over a purely static policy maker. However, it also shows that, even when considering dynamic impacts on the relevant market, a strong policy regime can still decrease the growth rate by decreasing incentives for low-productivity firms to enter into new markets. This is new to the literature consideration for the antitrust policy maker and something that warrants future study.

The chapter is organized as follows. Section 1.2 lays out the model and Section 1.3 details how the model is estimated. Section 2.4 examines how M&A affects growth at both the firm and aggregate levels. Finally, Section 1.5 examines the the relationships between antitrust policy, growth, and welfare.

1.1.1 Literature Review

Despite the significant concern surrounding the dynamic effects of antitrust policy, the literature is incredibly sparse and rarely considers the growth angle. The closest prior research to this paper in goal and approach is Cavenaile et al. [2021]. This paper used the Cavenaile et al. [2021] model as a base and then made several key advancements to increase its applicability. In the Cavenaile et al. [2021] paper, M&A opportunities arrive exogenously at a common rate for all pairs of firms in a market. The pair of potential merger partners then negotiate bilaterally. Their decision to have a rate be exogenous results in the limited ability of firms to respond to changes in antitrust policy. One advancement of the current paper is to replace the exogenous opportunities with endogenous search. The impact of an exogenous versus endogenous rate is most apparent when considering an antitrust policy shutdown, i.e., all M&A transactions occur freely. In this situation, Cavenaile et al. [2021]’s model demonstrates that there is only a -0.88% change in the growth rate and a minor impact on market concentration. When running the identical experiment,

this paper’s model results stand in stark contrast with a 10% increase in growth rate and a significant increase in the number of monopolies¹. These results have significantly different policy implications.

Another significant improvement that this paper makes is estimating the true function of the U.S. antitrust policy. Their paper makes the assumption that all transactions above the screening threshold² are obstructed with equal probability — an assumption that does not hold true in the empirical data. The consequence of their assumption is to underestimate antitrust enforcement for transactions that result in a high degree of market consolidation since transactions with extremely high HHIs and significant impact on market concentration would have the same probability of being blocked as a transaction that is just over the screening threshold and minimally impacts market concentration. This choice of model assumption would both bias their estimation (the parameter values are chosen to match the distribution of markups of the United States while allowing a higher frequency of large increases in markup transactions) as well as their result that shutting down antitrust enforcement has modest welfare consequences. Other studies have considered specific aspects of dynamic antitrust policy. [Fons-Rosen et al. \[2021\]](#) focuses on the impact of banning startup acquisitions and finds a positive growth effect. [Mermelstein et al. \[2020\]](#) studies the theory side of dynamic optimal antitrust policy.

¹The overall result of this increase in market concentration decreases welfare

²For the latest published guidelines, the screening threshold is an HHI of more than 0.15 and a change in HHI of more than 0.01.

A closely related strand of the literature focuses on the interaction of markup and innovation: [Peters \[2020\]](#); [Cavenaile et al. \[2021\]](#); [Phillips and Zhdanov \[2013\]](#); [Akçigit et al. \[2016\]](#); [Cavenaile et al. \[2021\]](#); and [Pearce and Wu \[2022\]](#). [Peters \[2020\]](#) shows that R&D activity can have both pro- and anti-competitive effects. Improving a firm’s existing product lines leads to increasing markups and misallocation while creating new products can displace high-markup incumbents. This paper’s model has a very similar finding to [Peters \[2020\]](#). [Cavenaile et al. \[2021\]](#) found that R&D provides firms with an increased ability to charge markups, resulting in static welfare losses for consumers. However, they found that the dynamic welfare gains were more significant than the static welfare loss. These findings echo the increasing need for a dynamic antitrust policy.

Additionally, there is a growing strand of the literature on the aggregate impacts of mergers and acquisitions without considering antitrust concerns. This include [Jovanovic and Rousseau \[2002\]](#), [Rhodes-Kropf and Robinson \[2008\]](#), [David \[2020\]](#), [Gaillard and Kankanamge \[2020\]](#), and [Bhandari et al. \[2023\]](#). These all show how mergers and acquisitions can improve the allocation of capital. However, by missing the potential for M&A to cause an increase in market power, it is difficult to determine the true welfare effects.

1.1.2 Institutional Details: U.S. Antitrust Policy

In the United States, antitrust authority is the dual responsibility of the Federal Trade Commission and the Department of Justice, Antitrust Division. While these are two distinct agencies that have different roles and responsibilities, this paper does not reach the level of detail where these differences are significant. As such, in this paper, the two agencies will be referred to as the antitrust authority. The statutory objective of the antitrust authority, found in the Clayton Act, is to prevent mergers where:

in any line of commerce or in any activity affecting commerce in any section of the country, the effect of such acquisition may be substantially to lessen competition or to tend to create a monopoly.

The authority's primary method of antitrust enforcement is the Premerger Notification Program, where a proposed M&A transaction³ is reviewed and, potentially, challenged by the government prior to the transaction⁴. This enforcement program is a process that has three primary steps:

1. The transacting parties inform the antitrust authorities of the proposed transaction.

³Only transactions over threshold prices are required to be reported in the Premerger Notification Program. This threshold changes by year. At the end of fiscal year 2019, the threshold value was \$90 million.

⁴While the government also conducts post-merger antitrust enforcement (for example, breaking up the Bell system) these actions are incredibly rare

2. The antitrust authority reviews the proposal, resulting in either a request for additional information, referred to as a second request, or approval for the transaction to be carried out.
3. For transactions that received a second request, the authority reviews the additional information and determines whether or not to challenge the transaction on antitrust grounds.

A challenge can attempt to block a transaction completely, but it also commonly requires the divestiture of a portion of one of the parties' firms. It is also worth noting that the merging parties can attempt to appeal the challenge in court.

A significant challenge when studying antitrust policy is a lack of transparent quantitative data surrounding the authority's decision-making process. The antitrust authority releases two quantitative resources. The first resource is the *Hart-Scott-Rodino Annual Report* (HSR report) in which the Federal Trade Commission and the Department of Justice Antitrust Division are required to report information related to that year's actions for the Premerger Notification Program. Quantitatively, it contains:

- the number of transactions reported to the antitrust regulator
- the number of requests for second reviews
- the number of transactions that the regulator attempted to block.

The HSR reports lack common concentration information, such as the Herfindahl-Hirschman Index, and only qualitatively describe the reasoning behind enforcement actions. Figure 1.2 plots the second request and challenge rate by time from fiscal year⁵ 1996 to fiscal year 2019. This graph shows that second requests are relatively rare and, over the plotted period, an average of 2.3% transactions received any form of challenge by the antitrust authority.

While systematic, quantitative data on the transactions that received second reviews is limited to the HSR reports, a FTC report *Horizontal Merger Investigation Data, Fiscal Years 1996-2011* allows a deeper look into the second review process. The first key data point from the report is the reason the antitrust authority initiated the second review – called the theory of potential violation. Of the identifiable reasons⁶, 80% of investigations theory of potential violation concerns horizontal mergers. This fraction is potentially a lower bound since other potential theories include "monopsony theory", "potential competition theory", and "joint venture" which could possibly result from horizontal mergers⁷. Given the antitrust authority's focus on horizontal mergers, this paper only includes horizontal mergers.

The other key data point is Table 3.1 of the report (which is reproduced in this paper

⁵The data contained within a year's report covers the U.S. Federal Government's fiscal year of Oct. 1 to Sept. 31.

⁶Non-identifiable reasons are "filing withdrawn by parties during the investigation", "closed after a quick look", and "investigations still open".

⁷All the theory of potential violation other than "vertical theory" comprise 92% of identifiable theory of potential violation

as Table 1.1). This provides the number of second reviews, as well as the number of enforcement actions, by the expected post-merger Herfindahl-Hirschman Index (HHI)⁸ and post-merger change in HHI. Panel A of Table 1.1 counts the number of second reviews. This shows the majority of the second reviews occur at higher levels of HHI and change in HHI. Panel B displays the fraction of second reviews resulting in an enforcement action. Generally, the higher the post-merger HHI and the higher the change in HHI, the higher the likelihood of an enforcement action. However, in the data, this likelihood is not monotonically increasing⁹.

1.2 Model

1.2.1 Environment

Household: There is a single infinitely-lived representative household that discounts the future at the rate ρ . The household consumes a consumption good C_t , with a price P_t ; supplies one unit of labor to the market at a wage rate w_t ; and invests in

⁸The Herfindahl-Hirschman Index is a commonly used measure of concentration within a market. It is defined as the sum of each market participant's squared market share.

⁹For example, the cell for a 5000-6999 post-merger HHI and 1200-2499 change in HHI has a lower, observed, percentage of enforcement actions than the 4000-4999 post-merger HHI and 1200-2499 change in HHI cell and the 5000-6999 post-merger HHI and 800-1199 change in HHI

asset A_t , with a return r_t . Assets evolve as

$$\dot{A}_t = r_t A_t + w_t - P_t C_t.$$

The household determines its consumption, labor supply, and savings to maximize its discounted lifetime utility of:

$$U = \int_0^\infty e^{-\rho t} \log C_t dt. \quad (1.1)$$

Firms: The economy is populated by a measure m of multi-product firms indexed by f . The set of product markets in which a firm f competes at time t is denoted by $I_{f,t}$. Within each of these product markets $i \in I_{f,t}$, the firm produces a single good with good-specific productivity $q_{f,i,t}$. The good specific productivity is combined with a firm-level productivity z_f . This productivity is constant across time. The firm f 's value at time t is denoted by $V_t(z_f, \{q_{f,i,t}\}; \Omega_t)$ where Ω_t is the states of all firms and markets at time t . For increased clarity, the full state of the firm will be frequently omitted with a " \rightarrow " symbol denoting changes to a firm's state

Product Markets: Differentiated product markets are indexed by $i \in [0, 1]$. For each product market i , there is a finite set of strategic firms $F_{i,t}$ and a competitive fringe, each producing differentiated varieties of products. For each firm $f \in F_{i,t}$,

their output is denoted $y_{f,i,t}$. The fringe's output is $y_{c,i,t}$. Within a product market i , the firms' goods are aggregated together into a product market output $y_{i,t}$ by a competitive firm employing C.E.S. technology with elasticity of substitution θ :

$$y_{i,t}^{\frac{\theta-1}{\theta}} = \sum_{f \in F_{i,t}} y_{f,i,t}^{\frac{\theta-1}{\theta}} + y_{c,i,t}^{\frac{\theta-1}{\theta}}. \quad (1.2)$$

Final Goods: All product markets' goods $y_{i,t}$ are aggregated into a final good Y_t by a competitive final goods firm. This firm employs a Cobb-Douglas aggregator¹⁰:

$$\log Y_t = \int_0^1 \log y_{i,t} di \quad (1.3)$$

Production and Competition: Within each market $i \in I_{f,t}$ in which the firm operates, they combine their common productivity z_f , their good specific productivity $q_{f,i,t}$, and a labor input $h_{f,i,t}$ to produce output $y_{f,i,t}$ by

$$y_{f,i,t} = z_f q_{f,i,t} h_{f,i,t}. \quad (1.4)$$

¹⁰In general, this technology could easily be replaced by a general CES aggregator; however, the unitary elasticity aggregator adds significant clarity during policy analysis.

The competitive fringe hire labor $h_{c,i,t}$ to produce output $y_{c,i,t}$ by

$$y_{c,i,t} = q_{c,i,t} h_{c,i,t}. \quad (1.5)$$

Within each market, the firms in $F_{i,t}$ compete in Cournot competition. For firm f in market $i \in I_{f,t}$, the firm, taking their competitors' prices as given, picks their price, output, and labor input to solve

$$\pi_{f,i,t} = \max_{p_{i,t}, p_{f,i,t}, y_{f,i,t}, h_{f,i,t}} p_{f,i,t} y_{f,i,t} - w_t h_{f,i,t} \quad (1.6)$$

$$s.t. \ y_{f,i,t} = z_f q_{f,i,t} h_{f,i,t} \quad (1.7)$$

$$y_{f,i,t} = p_{f,i,t}^{-\theta} p_{i,t}^{\theta-\eta} Y_t \quad (1.8)$$

$$p_{i,t}^{1-\theta} = \sum_{f' \in F_{i,t}} p_{f',i,t}^{1-\theta} + p_{c,i,t}^{1-\theta} \quad (1.9)$$

$$y_{f,i,t}, h_{f,i,t} \geq 0. \quad (1.10)$$

Research and Development: The firms conduct research and development to achieve two goals: (1) to improve the productivity of their current products and (2) to expand into new markets by creating new products. Each good's productivity sits on a quality ladder with step size $\lambda > 1$. Each time a firm improves its product, the good's productivity takes a single step up the ladder — its productivity goes from $q_{f,i,t}$ to $\lambda q_{f,i,t}$. The resulting change in firm f 's value from an improvement in its

market $i \in I_{f,t}$ product is

$$V(q_{f,i,t} \rightarrow \lambda q_{f,i,t}) - V.$$

Improvements arrive to the firm exogenously according to a Poisson arrival process at an endogenous innovation rate $j_{f,i,t}$. The cost of achieving that rate is $c_j(j_{f,i,t}) Q_t$ units of goods where c_j is assumed to be strictly convex and differentiable and Q_t is the aggregate productivity index. The improvements are market-specific. The firm can simultaneously attempt to improve all its products. Firm f chooses its innovation rate $j_{f,i,t}$ in market $i \in I_{f,t}$ to solve

$$\max_{j_{f,i,t} \geq 0} \{j_{f,i,t} (V(q_{f,i,t} \rightarrow \lambda q_{f,i,t}) - V) - c_j(j_{f,i,t}) Q_t\}. \quad (1.11)$$

A product's productivity never decreases; however, once a firm's product's productivity is \bar{n} steps behind the market leader, its product becomes defunct.

Creating a new product through expansion results in the firm expanding into a new market. This research effort is undirected, meaning the firm cannot target which market to enter. When they develop a new product, the firm receives a uniform random draw of which market $i \in [0, 1]$ it enters. Upon entering the new market, the productivity of the new good is a step above the maximum good-specific productivity currently existing in that market. Let $q_{\ell,i,t} = \max\{q_{f,i,t} : f \in F_{i,t}\}$. The new product's productivity is $\lambda q_{\ell,i,t}$. This is intended to capture new products

as superior versions of existing products. The expected value to the firm of a new product is

$$\mathbb{E}_{i'} \left[V \left(\{q_{f,i,t}\}_{i \in I_{f,t}} \rightarrow \{q_{f,i,t}\}_{i \in I_{f,t}} \cup \{q_{f,i',t}\} \right) \right] - V.$$

New products arrive to the firm according to a Poisson arrival process at an endogenous expansion rate $x_{f,t}$. The cost of achieving that rate is $c_x(x_{f,t}) Q_t$ units of goods where c_j is assumed to be strictly convex and differentiable. The firm chooses its expansion rate to satisfy

$$\max_{x_{f,t} \geq 0} \left\{ x_{f,t} \left(\mathbb{E}_{i'} \left[V \left(\{q_{f,i,t}\}_{i \in I_{f,t}} \rightarrow \{q_{f,i,t}\}_{i \in I_{f,t}} \cup \{q_{f,i',t}\} \right) \right] - V \right) - c_x(x_{f,t}) Q_t \right\}. \quad (1.12)$$

Within each product market, there is a maximum size cap of \bar{F} firms. If the market is drawn, and it already has \bar{F} firms, then the firm with the least profitable product¹¹, in terms of sales minus wages paid to production workers.

The competitive fringe does not perform R&D. Instead, their productivity is a fraction η of the market leaders's good's productivity. That is

$$q_{c,i,t} = \eta q_{i,t} \quad (1.13)$$

¹¹If multiple firms are tied for least profitable, one is picked with equal probability to exit the market.

where

$$q_{i,t} = \max\{q_{f,i,t} : f \in F_{i,t}\}. \quad (1.14)$$

Mergers and Acquisitions: The second source of firm growth is through mergers and acquisitions. This chapter focuses on horizontal mergers meaning M&A transactions are between current market incumbents and for a single market's product. While M&A transactions are traditionally thought of as one company buying the entirety of another, it is also frequently observed in the data that one company buys only a portion of another¹². Within market i , firm $f \in F_{i,t}$ receives an opportunity to attempt to purchase firm $f' \in F_{i,t}$ market i product at rate $a_{f',f,i,t}$. The potential buyer endogenously sets this rate $a_{f',f,i,t}$ at a cost $c_a(a_{f',f,i,t})Y_t$. When the parties agree to undertake a transaction, the prices are set by Nash bargaining, where the buyer's bargaining power is β . After the transaction is complete, the selling party exits the market. Crucially, from a competition perspective, that means there is one less competitor in the market.

After the buyer acquires the seller's product, there are two options. The first option is the buyer can either combine its product with the seller's product to produce a single, potentially improved product. The second option is for the buyer to shelve the seller's product, effectively killing it. This second option results in no change to the buyer's productivity within the market. Combining the two products is not

¹²For example, Google purchased only Motorola's cell phone division.

free. After the seller's product is transferred to the buyer, the buyer observes the combination cost $\delta_{f',f,i,t}$, which is drawn from an exponential distribution with shape parameter ω . If the buyer chooses to combine the products, the resulting productivity $\tilde{q}_{f,i,t}$ is

$$\tilde{q}_{f,i,t} = \zeta q_{f,i,t}^\alpha q_{f',i,t}^{1-\alpha}. \quad (1.15)$$

If the buyer chooses to shelve the seller's product, then the buyer's product's productivity remains unchanged.

Prior to every M&A transaction, the antitrust authority reviews the transaction and can choose to block the transaction. In the baseline model, the antitrust authority only considers the post-merger HHI and the change in HHI¹³. Let $\text{hhi}_{f',f,i,t}$ denote the post-merger HHI of firm f buying the market i good of firm f' at time t and $\Delta\text{hhi}_{f',f,i,t}$ denote the resulting change in HHI. The antitrust authority blocks a transaction with probability $B_{f',f,i,t} = B(\text{hhi}_{f',f,i,t}, \Delta\text{hhi}_{f',f,i,t})$. This function is an input to the model and is estimated in Section 1.3.1. If the regulator blocks a transaction, the parties do not have a chance to restructure the transaction and the deal is dead.

The total surplus for firm f buying the good of firm f' in market i at time t and

¹³This assumption allows the model to take full advantage of the limited, publicly-available data described in Section 1.1.2.

combining the goods, denoted $TS_{f',f,i,t}^C$, is

$$\begin{aligned}
TS_{f',f,i,t}^C &= V_{f,t} \left((q_{f,i,t}, q_{f',i,t}) \rightarrow (\zeta q_{f,i,t}^\alpha q_{f',i,t}^{1-\alpha}, 0) \right) \\
&\quad + V_{f',t} \left((q_{f,i,t}, q_{f',i,t}) \rightarrow (\zeta q_{f,i,t}^\alpha q_{f',i,t}^{1-\alpha}, 0) \right) \\
&\quad - V_{f,t} - V_{f',t} - \varepsilon_{f',f,i,t}.
\end{aligned} \tag{1.16}$$

The total surplus for the same transaction where the buyer kills the product, denoted $TS_{f',f,i,t}^K$, is

$$TS_{f',f,i,t}^K = V_{f,t}(q_{f',i,t} \rightarrow 0) + V_{f',t}(q_{f',i,t} \rightarrow 0) - V_{f,t} - V_{f',t}. \tag{1.17}$$

Given the Nash bargaining structure of M&A transactions, the buyer's surpluses are defined as

$$BS_{f',f,i,t}^C = \beta TS_{f',f,i,t}^C \tag{1.18}$$

$$BS_{f',f,i,t}^K = \beta TS_{f',f,i,t}^K. \tag{1.19}$$

Define firm f 's gain from combining $\Delta BS_{f',f,i,t}^C$ as

$$\Delta BS_{f',f,i,t}^C = BS_{f',f,i,t}^C - BS_{f',f,i,t}^K. \tag{1.20}$$

The probability of the firm combining $P_{f',f,i,t}^C$ is

$$P_{f',f,i,t}^C = 1 - \exp \left(-\omega \max \{0, \Delta BS_{f',f,i,t}^C\} \right). \quad (1.21)$$

Thus the expected buyer's surplus $\mathbb{E}[BS_{f',f,i,t}]$ is

$$\mathbb{E}[BS_{f',f,i,t}] = \mathbb{E} \left[(1 - B_{f',f,i,t}) (P_{f',f,i,t}^C BS_{f',f,i,t}^C + (1 - P_{f',f,i,t}^C) BS_{f',f,i,t}^K) \right]. \quad (1.22)$$

When the buyer decides its $a_{f',f,i,t}$, it solves

$$\max_{a_{f',f,i,t}} \{a_{f',f,i,t} \mathbb{E}[BS_{f',f,i,t}] - c_A(a_{f',f,i,t}) Y_t\}. \quad (1.23)$$

To simplify notation, let $MA_{f',f,i,t}(\varepsilon_{f',f,i,t})$ denote the change of state when firm f buys the good of firm f' in market i where the combination cost is $\varepsilon_{f',f,i,t}$.

Firm's Value Function: Combing the previous description results in the definition

of the firm's value function of

$$\begin{aligned}
r_t V_{f,t} - \dot{V}_{f,t} = & \underbrace{\sum_{i \in I_{f,t}} \pi_{f,i,t}}_{\text{Current Gross Profits}} \\
& + \underbrace{\sum_{i \in I_{f,t}} \max_{j_{f,i,t}} \{j_{f,i,t} (V_{f,t} (q_{f,i,t} \rightarrow \lambda q_{f,i,t}) - V_{f,t}) - c_J(j_{f,i,t}) Y_t\}}_{\text{Firm's R\&D Improvement}} \\
& + \underbrace{\sum_{i \in I_{f,t}} \sum_{\substack{f' \in F_{i,t} \\ f' \neq f}} j_{f',i,t} (V_{f,t} (q_{f',i,t} \rightarrow \lambda q_{f',i,t}) - V_{f,t})}_{\text{Competitors' R\&D Improvement}} \\
& + \underbrace{\max_{x_{f,t}} \{x_{f,t} \mathbb{E}_i D_{f,i,t} - c_x(x_{f,t}) Y_t\}}_{\text{Entry}} \\
& + \underbrace{\sum_{i \in I_{f,t}} \sum_z x_{z,t} (V_{f,t} (\{q_{f',i,t}\} \rightarrow \{q_{f',i,t}\} \cup q_{f'',i,t}) - V_{f,t})}_{\text{New Competitor's Entry}} \\
& + \underbrace{\sum_{i \in I_{f,t}} \sum_{\substack{f' \in F_{i,t} \\ f' \neq f}} \max_{a_{f',f,i,t}} \{a_{f',f,i,t} \mathbb{E} [BS_{f',f,i,t}] - c_A(a_{f',f,i,t}) Y_t\}}_{\text{Own M\&A Actions}} \\
& + \underbrace{\sum_{i \in I_{f,t}} \sum_{\substack{f', f'' \in F_{i,t} \\ f' \neq f''}} a_{f'',f',i,t} ((1 - B_{f'',f',i,t}) \mathbb{E} [V_{f,t} (MA_{f'',f',i,t} (\varepsilon_{f'',f',i,t}))]) - V_{f,t}}_{\text{Competitors' M\&A Actions}}
\end{aligned} \tag{1.24}$$

While this value function is incredibly complicated, it can be solved in parts according to the following lemma.

Lemma 1. *The firm's value function $V_{f,t}$ can be written as*

$$V_{f,t} = HQ_{f,t} + \sum_{i \in I_{f,t}} D_{f,i,t} \quad (1.25)$$

where $HQ_{f,t}$ is defined by

$$r_t HQ_{f,t} - \dot{H}Q_{f,t} = \max_{x_{f,t}} \{x_{f,t} \mathbb{E}_i D_{f,i,t} - c_x(x_{f,t}) Y_t\} \quad (1.26)$$

and $D_{f,i,t}$ is defined by

$$\begin{aligned}
r_t D_{f,i,t} - \dot{D}_{f,i,t} = & \underbrace{\pi_{f,i,t}}_{\text{Current Gross Profits}} \\
& + \underbrace{\max_{j_{f,i,t}} \{j_{f,i,t} (D_{f,i,t} (q_{f,i,t} \rightarrow \lambda q_{f,i,t}) - D_{f,i,t}) - c_J(j_{f,i,t}) Y_t\}}_{\text{Firm's R\&D Improvement}} \\
& + \underbrace{\sum_{\substack{f' \in F_{i,t} \\ f' \neq f}} j_{f',i,t} (D_{f,i,t} (q_{f',i,t} \rightarrow \lambda q_{f',i,t}) - D_{f,i,t})}_{\text{Competitors' R\&D Improvement}} \\
& + \underbrace{\sum_z x_{z,t} (D_{f,i,t} (\{q_{f',i,t}\} \rightarrow \{q_{f',i,t}\} \cup q_{f'',i,t}) - D_{f,i,t})}_{\text{New Competitor's Entry}} \\
& + \underbrace{\sum_{\substack{f' \in F_{i,t} \\ f' \neq f}} \max_{a_{f',f,i,t}} \{a_{f',f,i,t} \mathbb{E}[BS_{f',f,i,t}] - c_A(a_{f',f,i,t}) Y_t\}}_{\text{Own M\&A Actions}} \\
& + \underbrace{\sum_{\substack{f', f'' \in F_{i,t} \\ f', f'' \neq f \\ f' \neq f''}} a_{f'',f',i,t} ((1 - B_{f'',f',i,t}) \mathbb{E}[D_{f,i,t} (MA_{f'',f',i,t}(\varepsilon_{f'',f',i,t}))] - D_{f,i,t})}_{\text{Competitors' M\&A Actions}}
\end{aligned} \tag{1.27}$$

Proof. The proof is an accounting exercise of applying (1.25) to (1.24). ■

Distribution: A market i 's state at time t is the firms' productivity and goods' relative productivity¹⁴ $S_{i,t} = \{(z_1, \tilde{q}_1), \dots, (z_n, \tilde{q}_n)\}$. Let $\mu_t(S_t)$ denote the measure

¹⁴Relative productivity is the firm's good's productivity relative to the market's maximum good's productivity.

of markets in state S_t . This is defined by the Kolmogorov forward equation for markets

$$\begin{aligned}
\dot{\mu}_t(S_t) = & - \sum_{k=1}^{n(S_t)} j_k(S_t) \mu_t(S_t) \\
& + \int_{\{S'_t: \text{Firm } k \text{ innovation means } S'_t \text{ becomes } S_t\}} j_k(S'_t) d\mu_t(S'_t) \\
& - \sum_z x_{z,t} \mu_t(S_t) \\
& + \int_{\{S'_t: \text{Entry of firm type } z \text{ means } S'_t \text{ becomes } S_t\}} x_{z,t} d\mu_t(S'_t) \\
& - \sum_{k=1}^{n(S_t)} \sum_{k' \neq k} a_{k',k}(S_t) \mu_t(S_t) \\
& + \int_{\{S'_t: \text{Firm } k' \text{ and } k \text{ merge means } S'_t \text{ becomes } S_t\}} (1 - B_{k',k,t}(S'_t)) a_{k',k,t}(S'_t) d\mu_t(S'_t).
\end{aligned} \tag{1.28}$$

The distribution of divisions comes directly from the distribution of markets.

Market Clearing: Final goods are split among consumption and investment by

$$Y_t = C_t + \int_0^1 \sum_{f \in F_{i,t}} c_J(j_{f,i,t}) Y_t di + \int_0^m c_X(x_{f,t}) Y_t df. \tag{1.29}$$

The aggregate entry rate $X_{z,t}$ must match the aggregated individual expansion rates

$$X_{z,t} = \int_{\{f: z_f = z\}} x_{f,t} df. \tag{1.30}$$

Within the labor market, markets clear when aggregate labor demand equals one.

That is

$$1 = \int_0^1 \sum_{f \in F_{i,t}} h_{f,i,t} + h_{c,i,t} \, di. \quad (1.31)$$

Finally, household assets must equal the value of the economy's firms:

$$A_t = \int_0^m V_{f,t} \, df \quad (1.32)$$

1.2.2 Equilibrium

This paper focuses on the impact of policy on the balanced growth path.

Definition 2. *A **balanced growth path** is a set of firm-level allocations*

$$\left\{ x_{f,t}, \left\{ y_{f,i,t}, h_{f,i,t}, j_{f,i,t}, \{ a_{f',f,i,t} \}_{f' \in F_{i,t}, f' \neq f} \right\}, y_{c,i,t}, h_{c,i,t} \right\}_{i,t},$$

market distribution μ_t , aggregate allocations $\{C_t, A_t, Y_t\}_t$, entry rate $X_{z,t}$, and prices

$\left\{ P_t, r_t, w_t, \{ p_{f,i,t} \}_{f,i,t} \right\}$ such that

1. Consumption C_t and assets A_t maximize household welfare (3.8).
2. The firm allocation solve the optimization problems in (1.26) and (1.27).
3. The distribution of markets μ_t satisfies (1.28) and $\dot{\mu}_t = 0$.
4. The market clearing conditions (1.29)-(1.32) all hold.

5. *Aggregate output, consumption, wages, investment, and productivity all grow at the same rate g .*

Part 5 of the definition implies

$$Y_t = Y e^{gt} \tag{1.33}$$

$$C_t = C e^{gt} \tag{1.34}$$

$$w_t = w e^{gt} \tag{1.35}$$

$$Q_t = Q e^{gt}. \tag{1.36}$$

1.2.3 Equilibrium Results

On the balanced growth path, several analytic results can be derived.

Interest Rates: The interest rate is given by

$$r_t = \rho + g. \tag{1.37}$$

Welfare: On the balanced growth path, household welfare is

$$U = \frac{1}{\rho} \log C + \frac{1}{\rho^2} g.$$

Prices and Production: Given that within-market competition is assumed to follow the Cournot game, prices are pinned down via a nonlinear system

$$s_{f,i,t} = \frac{p_{f,i,t}^{1-\theta}}{\sum_{f' \in F_{i,t}} p_{f',i,t}^{1-\theta} + p_{c,i,t}^{1-\theta}} \quad (1.38)$$

$$\varepsilon_{f,i,t} = \theta + (1 - \theta)s_{f,i,t} \quad (1.39)$$

$$p_{f,i,t} = \frac{\varepsilon_{f,i,t}}{\varepsilon_{f,i,t} - 1} \frac{w_t}{z_{f,t} q_{f,i,t}} \quad (1.40)$$

where $s_{f,i,t}$ is firm f 's market share in market i and $\varepsilon_{f,i,t}$ is firm f 's market i elasticity. After computing prices, the output is determined by (1.8), and the labor input is determined by (1.4). To simplify things going forward, define

$$\begin{aligned} q_{i,t} &= \max\{q_{f,i,t} : f \in F_{i,t}\} \\ z_{i,t} &= \max\{z_{f,t} : f \in F_{i,t}\} \\ \tilde{q}_{f,i,t} &= \frac{q_{f,i,t}}{q_{i,t}} \\ \tilde{z}_{f,i,t} &= \frac{z_{f,t}}{z_{i,t}} \\ m_{f,i,t} &= \frac{\varepsilon_{f,i,t}}{1 - \varepsilon_{f,i,t}} \\ m_{i,t} &= \left(\sum_{f \in F_{i,t}} \left(\frac{m_{f,i,t}}{\tilde{z}_{f,i,t} \tilde{q}_{f,i,t}} \right)^{1-\theta} + \left(\frac{\eta}{z_{i,t}} \right)^{1-\theta} \right)^{\frac{1}{1-\theta}}. \end{aligned}$$

Aggregate Output: Aggregate output can be expressed as

$$Y_t = Q_t Z_t \mathcal{M}_t \quad (1.41)$$

where the goods' productivity index Q_t is defined by

$$\log Q_t = \int_0^1 \log q_{i,t} \, di, \quad (1.42)$$

the firms' productivity index Z_t is defined by

$$\log Z_t = \int_0^1 \log z_{i,t} \, di, \quad (1.43)$$

and a mark up term \mathcal{M}_t defined by

$$\mathcal{M}_t = \frac{\exp \left(\int_0^1 \log m_{i,t}^{-1} \, di \right)}{\int_0^1 \left(\sum_{f \in F_{i,t}} m_{f,i,t}^{-\theta} \tilde{z}_{f,i,t}^{\theta-1} \tilde{q}_{f,i,t}^{\theta-1} + \left(\frac{\eta}{z_{i,t}} \right)^{\theta-1} \right) m_{i,t}^{\theta-1} \, di}. \quad (1.44)$$

Growth: On the balanced growth rate, the growth rate g is given by

$$\begin{aligned}
g = & \underbrace{\log(1 + \lambda)(J + X)}_{\text{Improvements and Entry}} \\
& + \underbrace{\int_0^1 \sum_{\substack{f', f \in F_{i,t} \\ f \neq f'}} a_{f', f, i, t} (1 - B_{f', f, i, t}) \mathbb{E} \left[\log \left(\frac{q_{\ell, f', f, i, t}^{MA}}{q_{\ell, i, t}} \right) \right] di}_{\text{M\&A}}. \tag{1.45}
\end{aligned}$$

where

$$J = \int_0^1 \sum_{\substack{f \in F_{i,t} \\ q_{f, i, t} = q_{i, t}}} j_{f, i, t} di \tag{1.46}$$

$$X = \sum_z X_{z, t} \tag{1.47}$$

and $q_{\ell, f', f, i, t}^{MA}$ is the market leader's good's productivity after firm f acquires the good of firm f' . For the M&A term, it is important to note that $\log \left(\frac{q_{\ell, f', f, i, t}^{MA}}{q_{\ell, i, t}} \right)$ can either be positive or negative. The M&A term also reveals the direct effect of antitrust policy on growth — blocking transactions where $\log \left(\frac{q_{\ell, f', f, i, t}^{MA}}{q_{i, t}} \right) > 0$ directly reduces the growth rate.

1.2.4 Numerical Solutions

The model does not admit an analytical solution and must be solved numerically. The state space is defined by the number of productivity states z , the maximum number

of firms in the market \bar{F} , and the maximum number of steps behind the market leader \bar{n} . The specific values are three productivity states, a maximum of five firms, and eight steps behind the leader. This results in approximately 50 thousand market states and approximately 225 thousand division states. Appendix B describes the numerical algorithm in detail.

1.3 Estimation

1.3.1 The Antitrust Policy Function

One of the crucial inputs to the model is the probability of antitrust enforcement $\mathbb{P}(B|hhi, \Delta hhi)$. Given the limited publicly available data on antitrust decision-making, estimating this function is challenging. An approach commonly taken by the existing literature¹⁵ is to assume all transactions with a post-merger HHI and change in HHI above the horizontal merger guidelines receive enforcement with a constant probability. However, as documented in Section 1.1.2, the probability of a transaction receiving an enforcement action increases with the post-merger HHI and change in HHI. This paper’s approach is to decompose the conditional probability function into two parts — one which can be directly estimated from the publicly available data and the other which will be estimated as part of the moment matching

¹⁵See Cavenaile et al. [2021].

exercise.

The law of total probabilities allows $\mathbb{P}(B|hhi, \Delta hhi)$ to be written as

$$\mathbb{P}(B|hhi, \Delta hhi) = \mathbb{P}(B|S, hhi, \Delta hhi) \mathbb{P}(S|hhi, \Delta hhi) + \mathbb{P}(B|\bar{S}, hhi, \Delta hhi) \mathbb{P}(\bar{S}|hhi, \Delta hhi). \quad (1.48)$$

Since the first step in the antitrust review process is a second review, $\mathbb{P}(B|\bar{S}, hhi, \Delta hhi) = 0$. Thus, it follows that (1.48) becomes

$$\mathbb{P}(B|hhi, \Delta hhi) = \mathbb{P}(B|S, hhi, \Delta hhi) \mathbb{P}(S|hhi, \Delta hhi). \quad (1.49)$$

This transforms the conditional probability of blocking into a combination of the conditional probability of a second review and the conditional probability of blocking given a second review. The remainder of this section will detail how each object is estimated.

Beginning with the conditional blocking probability given a second review $\mathbb{P}(B|S, hhi, \Delta hhi)$, Table 3.1 of *Horizontal Mergers Investigation Data, Fiscal Years 1996-2011* brakes

down the fraction of second reviews by U.S. antitrust authorities that received enforcement action by post-merger HHI and change in HHI. Performing logistic regression¹⁶ results in an estimate of $\mathbb{P}(B|S, \text{hhi}, \Delta\text{hhi})$ given by

$$\text{logit } \hat{\mathbb{P}}(B|S, \text{hhi}, \Delta\text{hhi}) = -0.21 + 1.86 \text{ hhi} + 0.1862 \Delta\text{hhi} + 10.34 \text{ hhi} \times \Delta\text{hhi}. \quad (1.50)$$

To evaluate the accuracy of this classifier, twenty percent of the data was withheld from training. For the test, a transaction is considered blocked if $\hat{\mathbb{P}}(B|S, \text{hhi}, \Delta\text{hhi}) > \frac{1}{2}$. The resulting precision¹⁷ is 0.80, and recall is 0.93. Thus, the estimate of the test data is relatively accurate.

The second term in (1.49), the conditional probability of a second review, can not be directly estimated from the data. While *Horizontal Mergers Investigation Data, Fiscal Years 1996-2011* does have the number of second reviews by post-merger HHI and change in HHI, there is no data source that provides the number of transactions that did not receive a second review by post-merger HHI and change in HHI. To alleviate this challenge, this paper adopts an approach that enables indirect estimation of a second review policy function. Consider the antitrust policymaker's decision problem to issue a request for a second review. Second reviews are not free for the antitrust policy maker: they take a substantial amount of staff work to review the

¹⁶More details on this regression are provided in Appendix C

¹⁷Precision and recall are common accuracy measures for classification methods. Here, precision is the fraction of transactions predicted to be blocked that were blocked in the test data. The recall is the fraction of transactions blocked in the test data predicted to be blocked.

information and issue a decision. Thus, an optimal antitrust policymaker only issues second-review requests to transactions that they believe they are likely to block after the second review.

The policymaker's second review problem can be modeled as a discrete choice problem. Out of the set of all M&A transactions, the policymaker needs to pick which set of transactions receives a second review subject to the budget it has allocated for antitrust enforcement. The antitrust authority orders transactions by blocking likelihood and picks transactions until it exhausts its budget. This results in a second review threshold \bar{s} where any transaction with

$$\hat{\mathbb{P}}(B|S, \text{hhi}, \Delta\text{hhi}) \geq \bar{s} \tag{1.51}$$

receives a second review, and every other transaction is passed unopposed. This threshold, \bar{s} , becomes a model parameter that will be indirectly estimated.

1.3.2 Functional Forms

The model admits arbitrary functional forms on the firm's cost functions. This paper will use simple, convex-cost functions

$$c_j(j_{f,i,t}) = \varphi_j j_{f,i,t}^{1+\gamma_j} \quad (1.52)$$

$$c_x(x_{f,t}) = \varphi_x x_{f,t}^{1+\gamma_x} \quad (1.53)$$

$$c_a(a_{f',f',i,t}) = \varphi_a a_{f',f',i,t}^{1+\gamma_a} \quad (1.54)$$

1.3.3 Parameter Estimation

In addition to second review outcome probability $B(\text{hhi}, \Delta\text{hhi}|S)$, the model has a total of seventeen parameters: α , β , η , γ_a , γ_j , γ_x , λ , m , ω , φ_a , φ_j , φ_x , θ , ζ , z_1 , z_3 , and \bar{s} . The discount rate ρ is set to a standard value of $\rho = 0.05$. Given the difficulty in credibly identifying R&D curvature parameters without exogenous R&D cost variation, this paper follows [Acemoglu et al. \[2018\]](#)'s suggestion of using $\gamma_x = \gamma_j = 1$. For a similar reason, this paper sets $\gamma_a = 1$. The remaining fourteen parameters are estimated using the method of simulated moments. The remaining portion of this subsection will describe the target moments and their data sources. The data underlying each moment is restricted to 1996-2019.

To capture the U.S.'s growth dynamics, the model is estimated to match the average National Income and Product Accounts' GDP growth rate of 2.41% as well as the

ratio of R&D investments to GDP of 0.04. Given the paper’s interest, it is critical to accurately depict the interaction of competition and research effort. For almost all parameter sets, this model will produce an inverted-U relationship between the level of competition and research intensity as documented by [Aghion et al. \[2005\]](#). The goal is to match the inverted-U shape in the data. Following [Cavenaile et al. \[2021\]](#), the model is estimated to match a quadratic regression of relative sales on the log of RD spending¹⁸. The data underlying this regression is Compustat firm-years with a positive R&D expenditure. Sales share is defined by the shares of sales within its four-digit SIC. The result of this regression is

$$\log \text{RD}_{f,t} = 1.39 \times \text{share}_{f,t} - 1.11 \times \text{share}_{f,t}^2 + \varepsilon_{f,t}. \quad (1.55)$$

The next set of moments are statistics on observed M&A transactions where the acquirer was a U.S. publicly traded corporation. The data source is SDC platinum, which contains records of individual M&A transactions. [Barnes et al. \[2014b\]](#) reports that it is generally a reliable and complete source of the target transactions. The first moment is the aggregate M&A transaction value to GDP of approximately 5% per year. Next is the mean and standard deviation of the merger premium. In order to discipline the killer acquisition rate, the model is targeted to match [Cunningham](#)

¹⁸In [Cavenaile et al. \[2021\]](#), they used patent counts and improvement arrival rates. Given the multi-product nature of this model, R&D spending adds up to a firm-level variable more easily than the arrival rates.

et al. [2021]’s lower bound on the estimated killer acquisition rate of 5.3%. The final M&A moment is the antitrust authority’s second request rate of 3%. The final set of moments is related to the firm’s productivity distribution.

These data moments are summarized in Panel B of Table 1.2. The parameters are chosen to match these target moments. The resulting parameters are shown in Panel A of Table 1.2 and the resulting model moments are shown in Panel B. While the model is not exactly identified, the simulated moments are close enough that the results of this parameter set should be reflective of the exactly identified model.

1.4 The Effect of M&A on Growth

The goal of this section is to understand the impact of M&A on R&D effort. The clearest way to examine the impact is to compare the model with an alternative where the antitrust regulator blocks all M&A. The next section examines the impact at the level of the individual firms. The following section examines the aggregate impact and the associated impact on growth.

1.4.1 Impact of M&A Across the Firm Distribution

To begin, consider the entry effort margin. Table 1.3 shows the firm-level entry effort in both the baseline model and the no M&A case by productivity level. Without

M&A, every firm decreases their entry effort showing that M&A and entry are complements. The 53% decrease for the low-productivity firms is the largest response. The reason is simple: removing M&A is removing a channel for firms to receive a return on their entry investment. While an individual, low productivity firm's entry effort is relatively small, they make up 50% of all firms. Thus, they comprise 36% of the aggregate entry rate¹⁹. Even the high productivity firm experiences a negative decline in their entry rate when there is no M&A. For the high productivity firms, the removal of the investment return channel of being bought out is part of the story. However, their profits are also affected by the reduced ability to consolidate markets²⁰.

After seeing how the entry margin of R&D decreased without the ability to perform M&A, a natural question is if the improvement margin is similarly impacted. Figure 1.3 displays the inverse cumulative distribution function of the change in improvement rate when M&A is removed²¹. It reveals that approximately 75% of the divisions would perform less improvement R&D in a world without M&A. This results in around a 19% decrease in the average improvement effort²². However, in addition to the changes in improvement rates, the distribution of divisions also changes

¹⁹The middle productivity firm comprises 46% of the aggregate entry rate, and the high productivity firm comprises the remaining 18%.

²⁰A simple measure of this is the number of monopoly and duopoly markets. In the baseline model, 24% of markets are monopolies or duopolies. Without M&A, it is nearly zero markets.

²¹The distribution of divisions underlying this figure is the distribution from the baseline model

²²Computing the average based on the baseline distribution of divisions.

between the two cases. The lack of M&A results in markets being more competitive. Given the model matches research and development's inverted U-shape relationship with competition²³, this increase in competitiveness results in a significant increase in average improvement effort compared to the baseline.

1.4.2 M&A and Aggregate Growth

The previous subsection reveals that M&A boosts entry effort and reduces improvement effort compared to a world without M&A. This subsection examines the aggregate impact. Aggregate growth can be decomposed into four components: improvement $g_{\text{Improvement}}$, entry g_{Entry} , M&A combination g_{Combine} , and M&A kill g_{Kill} . Table 1.4 decomposes the aggregate growth rate into these four components. The largest contributor, adding 1.29pp to the aggregate growth rate, is the entry component. Next is the improvement component adding 0.69pp. M&A resulting in combining products contributes 0.41pp. Somewhat surprisingly, M&A transactions resulting in the acquirer killing the acquired product only reduce the aggregate growth rate by 0.01pp.

To understand the impact of M&A on these aggregate components, consider the alternative model where all M&A is banned. The aggregate growth rate in this case is 1.9% which is 0.49pp lower than the baseline model. The decomposition

²³Documented in [Aghion et al. \[2005\]](#).

into the four components, as well as the difference between the baseline and the no M&A case, is also shown in Table 1.4. Consistent with the previous section, the improvement component does increase by 0.26pp. However, this does not make up for the -0.35pp change in the entry rate component and the -0.41pp change from producing new goods through M&A. This also reveals that the M&A substitutes with aggregate improvement R&D but is complementary with aggregate entry R&D. However, as the next section will show, the presence of M&A does increase aggregate misallocation.

1.5 Competition Policy

1.5.1 The Dual Edge Sword Effect of M&A

The previous section showed that banning all M&A would reduce the aggregate growth rate by nearly a half percentage point. Given that M&A both acts as an incentive for entry as well as a source of growth through the combining products, a reasonable hypothesis is that allowing all M&A transactions may increase growth. As Table 1.5 reveals, the growth rate is 0.24pp higher than the baseline case. However, it also reveals that the household would prefer not live in either the economy with no M&A or the economy with no enforcement. The difference in welfare from the baseline economy is determined by two things: the change in the economic growth

rate and the change in static misallocation. In the case where there is no M&A, the impact of the reduced misallocation does not make up for the reduced growth rate. In the case where there is no antitrust enforcement, the increased growth rate does not offset the increased static misallocation. The source of this increased misallocation is increased concentration in markets. For example, the fraction of monopoly markets rise from less than one percent in the no M&A case to approximately 51% in the no enforcement case²⁴. This increase in concentration reduces the level base level of consumption, C , by 24% compared to the baseline model. In both cases, the household's welfare decreases. The no M&A case has a decrease of 18.15% and the no enforcement case has a decrease of 7.86%. This shows that an optimal antitrust authority must consider its impact on growth in addition to its impact on static misallocation.

1.5.2 Simple Antitrust Rules

While the predictive analysis to identify anticompetitive mergers in the real world is extremely difficult, within the model, it is simple. Given the unitary elasticity final goods aggregator, market-level revenue $p_{i,t}y_{i,t}$ is equal across markets. If a merger results in a lessening of competition, it also results in a decrease in output. So, within the model, a simple rule could reject a proposed merger if it results in a decrease in

²⁴In the baseline model, approximately 4% of markets are monopolies.

some measure of output.

From a purely static perspective, a potential simple rule blocks the merger if the post-merger level of market output is less than the premerger level of market output. Let $\mathbb{E}y_{f',f,i,t}^MA$ be the expected level of output after the merger²⁵. The pre-merger level of output is $y_{i,t}$. So, the simple static rule is to block a merger if

$$\mathbb{E}y_{f',f,i,t}^MA < y_{i,t}. \quad (1.56)$$

The way to interpret this rule is that the antitrust policymaker disallows any merger that does not immediately increase a market's output.

The static nature of this rule means it ignores the possibility that output may increase in the market at some future date. For example, suppose two firms combine their products to produce a better product. The increase in concentration drops the level of sales competition. However, this better product may bring the remaining competitors closer to product quality parity. In response to this increase in quality, the firms increase their research efforts to improve their product. Thus, in the long run, the level of output may actually increase compared to if the M&A transaction did not occur.

This motivates the policymaker to consider the dynamic impact of the transaction.

²⁵The expectation operator is required since a merger can result in two possibilities: the acquiring combines the two firms' products or it kills the acquired companies products. Each case results in a different level of post-merger output.

A dynamic alternative to the simple static rule is to compare the discounted average future output paths. Let

$$E_{f',f,i,t}^{MA} = \mathbb{E} \int_0^\infty e^{-\rho s} y_{f',f,i,t+s}^{MA} ds \quad (1.57)$$

$$E_{i,t} = \mathbb{E} \int_0^\infty e^{-\rho s} y_{i,t+s} ds. \quad (1.58)$$

The equivalent dynamic rule to (1.56) is block a proposed merger if

$$E_{f',f,i,t}^{MA} < E_{i,t}. \quad (1.59)$$

The main difference between the two rules is that rule (1.59) accounts for the fact that the post-merger output may grow faster than the pre-merger path.

The results of these two proposed simple rules are shown in Table 1.6. The dynamic simple rule outperforms the static simple rule since the static simple rule has a significant impact on growth. This shows the importance for the policymaker to focus on the dynamic impacts of a M&A transaction on the market. However, the dynamic policy performs worse than the estimated baseline policy. This is because, while the rule takes into account the dynamic effects within the market, it does not account for the effect of the policy on a firm's entry effort decision. The entry growth component g_{entry} decreases by 0.1pp compared to the baseline. This shows that the policymaker needs to carefully consider the spillovers of its policy from individual

transactions to the firm's entry decisions.

1.5.3 Banning Killer Acquisitions

Cunningham et al. [2021] found, within the pharmaceutical industry, between 5.7% and 7.4% are killer acquisitions. This has raised alarms with antitrust policymakers. It is difficult to, ex-ante, determine which transaction will be killer acquisitions in data. However, in the model, the probability that a transaction results in a killer acquisitions is simple to compute. That also makes it straightforward to consider the impact of banning the transaction.

Consider one final antitrust policy rule. If the probability of a transaction being a killer acquisition is above 50%, the antitrust policymaker disallows that transaction. Otherwise, the antitrust policymaker follows the rule estimated in Section 1.3. This rule can be interpreted as adding an antitrust policy that bans transactions that are more likely than not to be killer transactions while keeping the remaining policy the same.

This rule results in a welfare increase of 0.3% and a decrease in the growth rate of 0.02pp²⁶. This small increase is due to killer acquisitions being a small fraction of overall acquisitions along with their minimal effect on growth²⁷. While banning

²⁶The reason for the decrease in the growth rate is once again a decrease in entry effort

²⁷In the baseline, killer acquisitions contribution to the growth rate was -0.01pp.

killer acquisitions is welfare improving, these results suggest that the antitrust authority would be better served spending its limited resources towards larger concerns. However, if the fraction of transactions that are killer acquisitions was to rise, this analysis would need to be revisited.

1.5.4 The Importance of Endogenous M&A for Policy Analysis

This chapter was the first where M&A rates in the model were chosen endogenously by the firms. The motivation to include this feature was to allow for the possibility that firms would ramp up their market consolidation in a less restrictive policy environment. Given that this chapter's policy analysis stands in stark contrast with the previous exogenous M&A literature²⁸, it is important to examine the impact this choice has on the results. To uncover the effect, consider an alternative version of the model where M&A rates are exogenous. For each division, fix the M&A arrival rate to the level in the baseline model. Rates vary across firm states²⁹ but are the same across different policy regimes.

When shutting down antitrust enforcement in the model with exogenous M&A rates,

²⁸For example [Cavenaile et al. \[2021\]](#) found virtually no difference, in terms of average markup or growth, between an economy with their estimated antitrust enforcement and one with no antitrust enforcement.

²⁹This is different from [Cavenaile et al. \[2021\]](#) where rates are exogenous and constant across firm states.

there is an approximately 18% increase in average markup (compared to 131% in the baseline model). Understanding this difference requires examining the firms' motives. While the lack of enforcement leads the potential for an increase in anti-competitive mergers, the firms can not respond to the change in policy. Their effort is determined under the baseline antitrust policy. Attempting M&A is costly meaning firms will minimize the effort they put into merger's that are likely to be blocked. Thus the arrival rate of significantly anti-competitive mergers is still relatively low. The result is that the average post-merger HHI goes from 2850 with enforcement to 3200 without enforcement. In the endogenous model, the average post-merger HHI is 4275. Thus this shows endogenous M&A is a crucial model feature.

1.6 Conclusion

This chapter contributes to the ongoing discussion surrounding growth and antitrust policy. It found that M&A's was a double-edged sword: both unrestricted and overall restrictive M&A cause decreased welfare. The novel model serves as a test bed for potential antitrust policy. It was found that dynamic policy substantially improved welfare compared to static policy. However, it exposed how sensitive economic growth can be to the antitrust policy regime. Specifically, it details a novel-to-the-literature impact of antitrust policy on less productive firms' entry decisions.

This chapter's contributions were just a first step. An open question worth future

study is how to balance the antitrust policymakers' goal of reducing misallocation within a market with the broader impact of the policy regime on entry. The second question is, given antitrust policy changes from one presidential administration to the next, how does aggregate policy uncertainty impact these results? It is possible that if agents expect a future, more lenient antitrust regime, the impact of more restrictive policies on growth may be muted.

1.7 Tables

Table 1.1: Outcome of Second Reviews by HHI and Δ HHI

Panel A. Number of Second Reviews

		Change in HHI (Δ HHI)							
		0-99	100-199	200-299	300-499	500-799	800-1199	1200-2499	2500+
Post-Merger HHI	0-1799	14	48	39	28	10	1		
	1800-1999	4	9	11	16	17			
	2000-2399	3	7	15	44	44	4		
	2400-2999	3	6	11	24	58	36		
	3000-3999	4	5	7	14	39	92		
	4000-4999		4	2	6	14	22	71	
	5000-6999	1	6	10	9	19	23	165	52
	7000+			1	1	3	9	27	248

Panel B. Percent of Second Reviews Resulting in a Challenge

		Change in HHI (Δ HHI)							
		0-99	100-199	200-299	300-499	500-799	800-1199	1200-2499	2500+
Post-Merger HHI	0-1799	0	35	49	61	30	0		
	1800-1999	0	56	45	75	71			
	2000-2399	33	14	47	57	73	50		
	2400-2999	33	67	55	75	76	72		
	3000-3999	25	60	71	64	64	77	74	
	4000-4999		50	50	83	71	82	96	
	5000-6999	100	100	80	89	100	91	88	90
	7000+			100	100	100	100	96	99

Note: The data presented in this table is a reproduction of *Horizontal Merger Investigation Data, Fiscal Years 1996-2011* produced by the Federal Trade Commission.

Table 1.2: Estimated Parameters and Target Moments

Panel A. Estimated Parameters

Parameter	Value	Parameter	Value
α	0.65	φ_j	0.22
β	0.56	φ_x	0.65
η	0.52	θ	6.54
λ	0.07	ζ	1.16
m	2.12	z_1	0.80
ω	26.58	z_3	1.27
φ_a	0.05	\bar{s}	0.82

Panel B. Target Moments

Description	Data	Model	Description	Data	Model
GDP Growth Rate	2.41	2.39	M&A-GDP Ratio	0.05	0.05
R&D-GDP Ratio	0.04	0.04	Avg. Merger Premium	41.56	42.33
Log Regression Sales	1.39	1.05	S.D. Merger Premium	44.90	43.54
Log Regression Sales ²	-1.11	-1.12	Killer Acquisition Rate	5.30	5.18
Avg. Markup	34.98	35.13	Second Request Rate	3.12	3.25
S.D. Markup	34.60	34.48	Bottom 50% Sales Share	1.40	1.23
Avg. Firm Profitability	0.15	0.15	Top 10% Sales Share	80.37	80.25

Table 1.3: Entry Effort By Firm Productivity

Productivity	Baseline	No M&A	Difference
z_1	0.15	0.07	-53%
z_2	0.24	0.21	-13%
z_3	0.36	0.33	-8%

Note: The column "Baseline" is the entry rate in the baseline model. The column "No M&A" is the entry where M&A is shut off. The "Difference" column is the difference between the no M&A case and the baseline model.

Table 1.4: Sources of Aggregate Growth (Percent)

Source	Baseline	No M&A	Difference
Total	2.39	1.90	-0.49
Improvement	0.69	0.95	0.26
Entry	1.29	0.94	-0.35
M&A Combine	0.41	0.00	-0.41
M&A Kill	-0.01	0.00	0.01

Note: The column "Baseline" is the decomposition of the aggregate sources of growth in the baseline model. The column "No M&A" is the decomposition where M&A is shut off. The "Difference" column is the difference between the no M&A case and the baseline model.

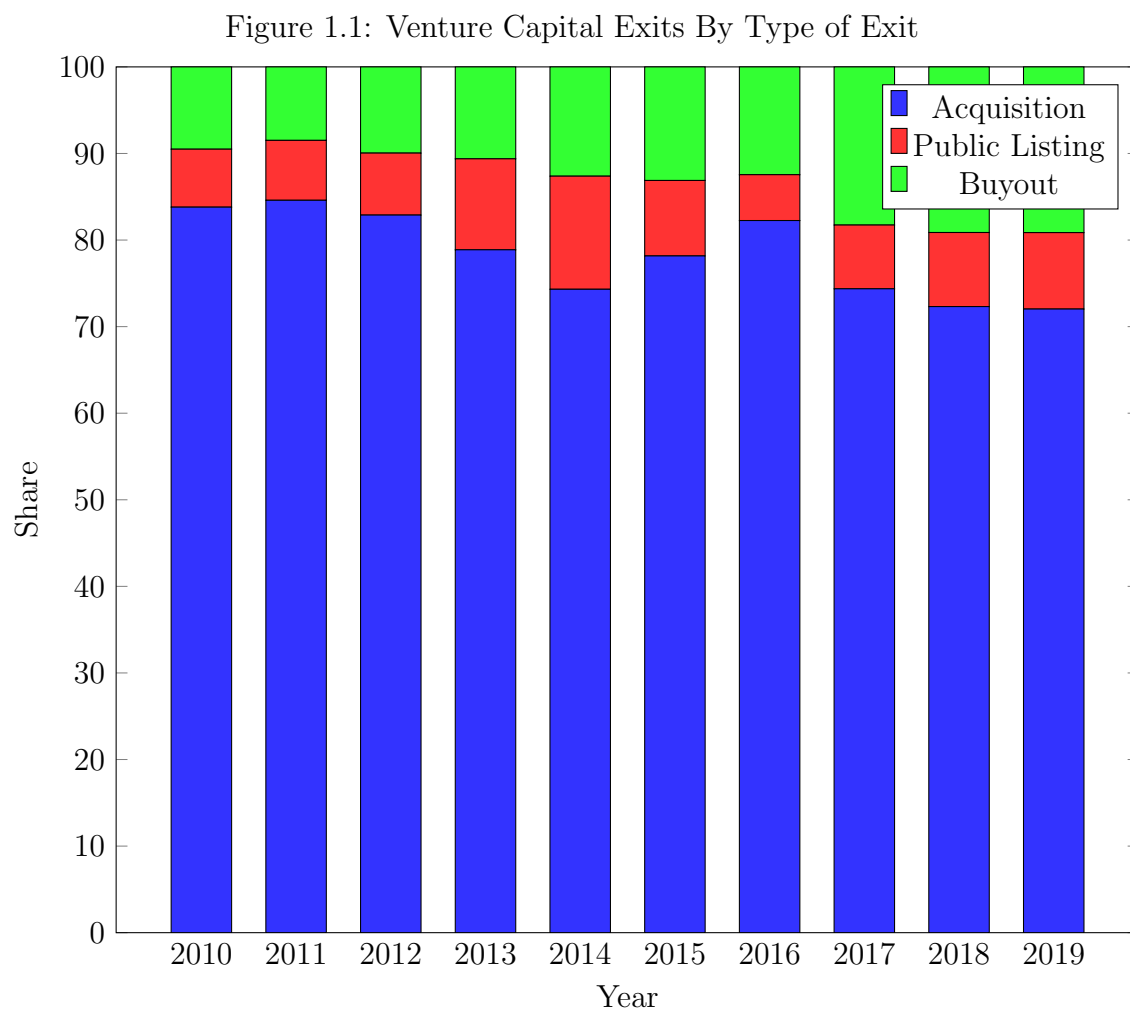
Table 1.5: Alternative Antitrust Policies

	Baseline	No M&A	No Enforcement
Welfare			
Welfare Difference (%)		-18.15	-7.86
C	0.96	0.99	0.73
g	2.39	1.90	2.63
Components of Y			
Z	1.04	1.01	1.08
\mathcal{M}	0.92	0.98	0.73
Sources of Misallocation			
Mean Markup(%)	35.13	23.32	81.28
S.D. Markup(%)	34.48	4.75	72.56

Table 1.6: Simple Antitrust Rules

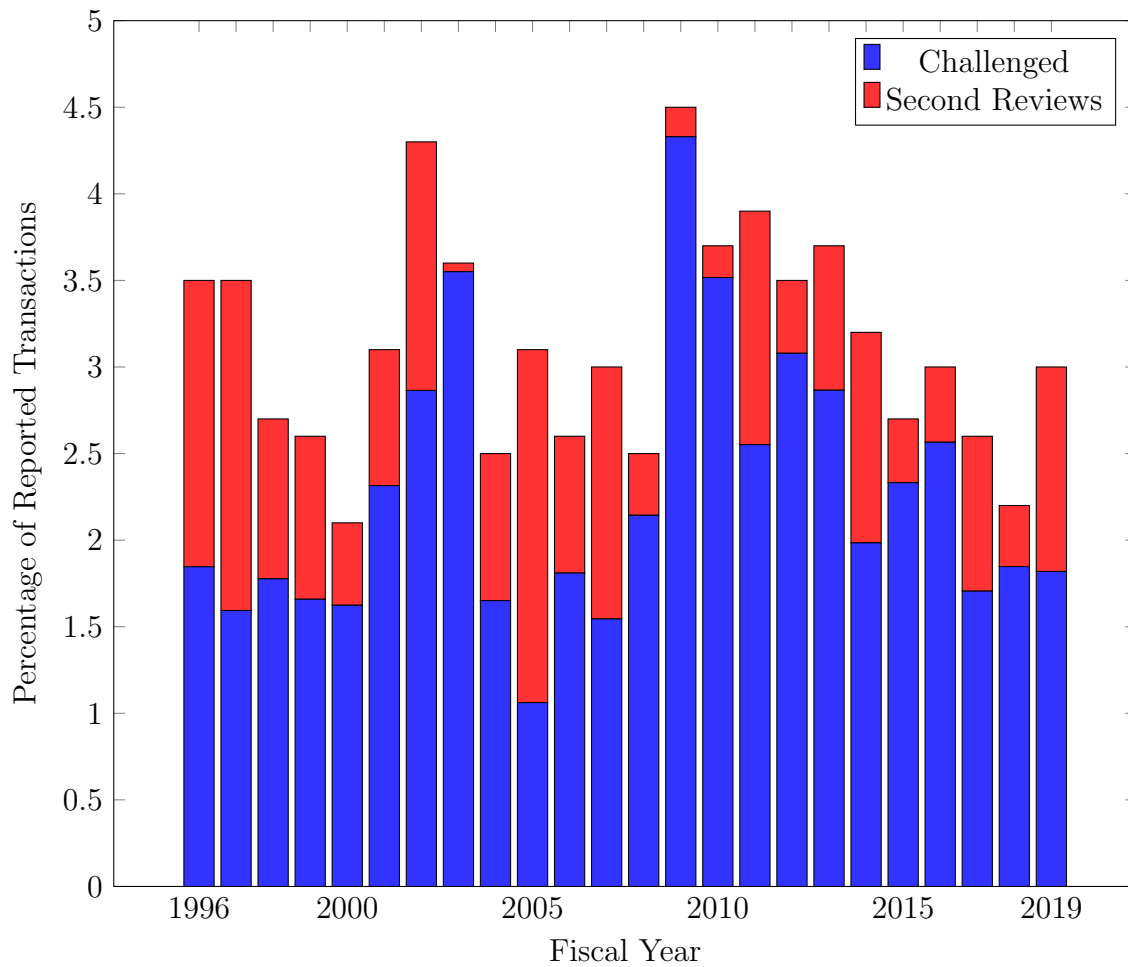
	Baseline	Static	Dynamic
Welfare			
Welfare Difference (%)		-9.09	-4.94
C	0.96	0.99	0.98
g	2.39	2.11	2.23
Components of Y			
Z	1.04	1.03	1.04
\mathcal{M}	0.92	0.96	0.94
Sources of Misallocation			
Mean Markup(%)	35.13	28.28	30.11
S.D. Markup(%)	34.48	13.22	23.86

1.8 Figures



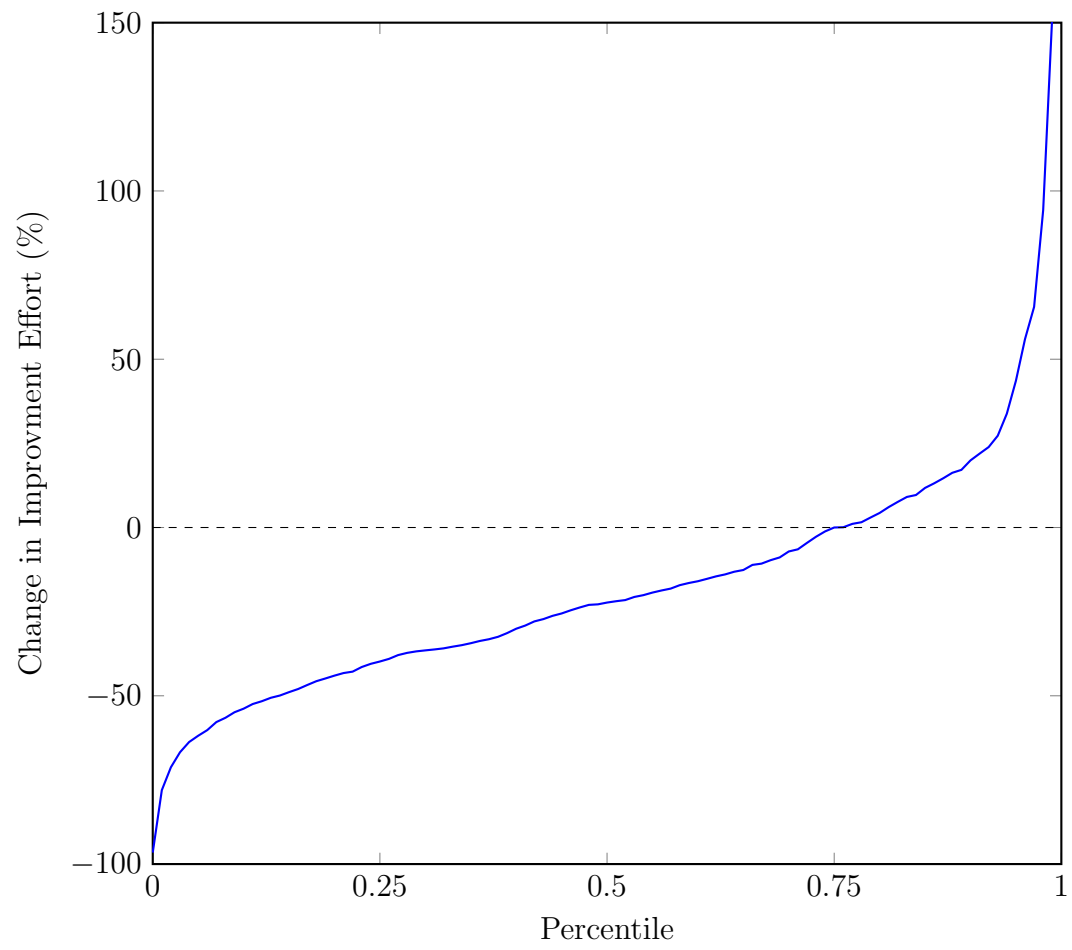
Note: This data was retrieved from the Q3 2022 Pitchbook-NVCA Venture Monitor.

Figure 1.2: Frequency of Second Antitrust Reviews and Challenges



Note: The percentage of transactions reported as part of the Premerger Notification Process which received second review and some form of challenge by the antitrust authority. The data is retrieved from the *Annual Reports to Congress Pursuant to the Hart-Scott-Rodino Antitrust Improvements Act of 1976* fiscal years 1996-2019. The report's time period is the government's Oct. 1st to Sept. 31st fiscal year instead of the calendar year.

Figure 1.3: Inverse Cumulative Distribution for Change in Improvement Effort



Chapter 2

Employing Natural Language

Processing To Understand M&A

Motives

2.1 Introduction

The previous chapter demonstrated the importance of merger and acquisition (M&A) to the overall macroeconomy. A remaining challenge is to understand the motivation for each M&A's transaction observed in data. A common motive in any M&A transaction is for the acquirer to increase their market power by purchasing their competitors and consolidating markets. This increase in market power can result

in economic misallocation. However, this is not the only possible motive. Observing the M&A history of tech companies, they frequently use M&A to expand their operations into new sectors¹. Such a transaction may have a different economic impact: the reallocation of assets from low-productivity firms to high-productivity firms. Reallocation can lead to economic gains, while an increase in misallocation leads to economic losses. These differences in motives are of key interest to antitrust policymakers. For M&A, the antitrust policymaker’s primary concern is limiting the ability of firms using M&A to increase their market power. However, policymakers must fight against concentration while limiting its impact on investment incentives. This chapter examines the M&A transactions involving U.S. companies. Between 1990 and 2019, there were 182 thousand such transactions. The goal is to distinguish if the transaction’s parties are horizontally and/or vertically linked. Given the number of transactions, it is infeasible for researchers to evaluate each transaction to determine the relationship between the transacting parties². If one were to hire human research assistants to examine each transaction, it would be an extremely expensive and time-consuming endeavor. Instead, this chapter uses a novel natural language approach employing large-language models (LLMs) to determine whether the transacting parties are close competitors.

The methodology compares textual descriptions of the acquired and acquiring parties

¹For example, Google used M&A to expand into the online mapping, self-driving cars, and mobile phone sectors. Sectors where they were not previously competing.

²This is also a challenge for the antitrust authority given their limited resources.

to determine if overlaps exist in their operations. While it uses machine learning, the steps are very similar to giving a task to a human research assistant. The machine learning tool is given a detailed set of instructions about its task to analyze M&A transactions as well as a copy of the current U.S. government’s merger guidelines. It is then given the textual description of both parties and asked to perform analysis. The result is a database of the linkage classifications for transactions in the sample that can be applied to any research task on M&A.

Similarly to a human research assistant, there are no guarantees of 100% accuracy and LLM’s analysis results need to be reviewed and verified. To evaluate the accuracy method, a subsample of transactions is drawn and manually labeled. After comparing the manual labeling to the LLM’s analysis, the LLM classifier is highly accurate when the descriptions have substantial information. However, for a number of transactions that have insufficient information, the classifier classifies with high confidence instead of responding that it is uncertain.

For the 95% of transactions that the model provides a definitive classification, the model showed that 39% of transactions are horizontally-linked only, 30% of transactions are vertically-linked only, 5% of transactions are both horizontally- and vertically-linked, and 17% of transactions are neither horizontally- nor vertically-linked. This pattern is relatively robust across industries and whether the parties are public companies or not. It is also robust from every presidential administration

in the data set.

The first paper to use textual descriptions to study M&A transactions was [Hoberg and Phillips \[2010\]](#). They used a simpler dictionary-based method to compare the words in M&A activity descriptions between public firms. This method measures the similarity of the two firms by how many common words they use in their descriptions. That original paper focused on horizontal integration. They then apply a similar methodology to vertical M&A in [Frésard et al. \[2019\]](#). One weakness of this method is that it is dependent on similar firms using similar wording, but if distinct synonyms are used to describe their business, the firms won't be considered similar. The method also requires the researchers to develop a notion of how to convert these vectors into measurements of the similarity of the two firms. They attempt to address this shortcoming in [Gerard and Phillips \[2021\]](#) by employing a machine-learning tool named doc2vec. While this helps to address the wording problem, it does not address requiring a general specification of similarity. This paper's approach uses the cutting-edge GPT large language model, which has a deeper understanding of the English language compared to doc2vec and does not require a similarity threshold to be defined. An additional consideration is that they only apply the doc2vec methodology on the 10-Ks of publicly traded firms. This will miss M&A activity involving a private firm. This paper instead uses the SDC platinum data set, which contains textual descriptions of private M&A transactions.

Large-language models³ have been transforming a number of fields of economic study. [Handlan \[2022\]](#) uses LLMs to study the impact of Federal Reserve communications on market expectations. [Kazinnik \[2023\]](#) employs a LLM to study responses to bank run scenarios. [Bajari et al. \[2023\]](#) employs an LLM to create hedonic price indices. This paper adds an additional application for large-language models with mergers and acquisitions data.

Section 2.2 describes the M&A transaction data. Section 2.3 then describes the prompt-based methodology to classify transactions as either vertically or horizontally linked. This methodology is then applied to the M&A transaction data in Section 2.4.

2.2 Data on U.S. M&A Transactions

The SDC Platinum data set (SDC), produced by Refinitiv, is the primary data source used for research into mergers and acquisition transactions. Refinitiv states that they use a combination of regulatory filings, press releases, and news reports. For transactions involving U.S. publicly traded firms, [Barnes et al. \[2014a\]](#) reports, compared to a hand collect set of M&A transactions, SDC is fairly complete from 1986 onward. The authors also report that "SDC is more prone to errors on smaller, high book-to-market acquirers with weak announcement period market responses".

³A large-language model is a neural network model for language comprehension trained on massive amounts of data. A complete introduction is beyond the scope of this article. For a helpful introduction, see [Ash and Hansen \[2023\]](#).

This suggests that SDC is a reliable source to study M&A transactions. In order to focus on U.S. M&A transactions, the base SDC data set is filtered on the following criteria:

1. The announcement date was after 1990
2. Both the acquiring and the target company is based in the U.S.
3. The ultimate parent for the acquirer and target are distinct⁴
4. Prior to the transaction, the acquirer had less than 50% of voting control over the target.
5. The acquirer was not a government entity⁵

These restrictions result in 237 thousand transactions. Table 2.1 contains summary statistics of the transactions. There are 65 thousand transactions with reported valuations. The mean transaction value is \$385 million while the median value is significantly less at \$35 million which reveals that the value distribution is significantly skewed to the right. SDC Platinum reports Standard Industry Classification (SIC) codes for the target firms⁶. In terms of valuation, the majority of target firms are in

⁴Sometimes corporate reorganizations enter the data set as a transaction for regulatory purposes. These are not true M&A transactions since they are acquirer and target are under common control.

⁵For example, the U.S. government's 2008 takeover of Freddie Mac and Fannie Mac are included in the data set. Also, several private hospitals are being bought by the state or local government. While these transactions are interesting, they are not the focus of this paper and thus are excluded.

⁶SDC Platinum does not report the SIC code for the acquirer

finance, tech, communications, and pharmaceuticals.

The majority of transactions involve publicly traded companies⁷. When both sides of the transaction are public, the valuations are higher, but the industry composition remains roughly the same. When a transaction involves a public company that buys a private company, the valuations are smaller, and target companies are more concentrated in the tech sector. Approximately 22% of transactions in the data set are between private companies. These transactions are generally of smaller valuations and less tech-focused and pharmaceutical-focused.

Turning to trends over time, since the year 2000, the number of transactions has been relatively steady, around six thousand transactions a year, while the median transaction value has been increasing. Panel A of Figure 2.1 shows the number of transactions per year. The number of transactions fell during the Dot-Com and the Global Financial Crises but quickly rebounded. Panel B of the figure shows the median value of a M&A transaction has increased from an average of \$26 million prior to 2000 to \$53 million after 2010.

Over this period, the share of transactions between both parties, which are public companies, has been decreasing. Figure 2.2, breaks down transactions involving publicly traded corporations into shares where both parties were public, where only the acquiring company is public, and where only the target company is public. During

⁷This could be the total acquisition of a public company or a transaction for only one of its subsidiaries

the 1990s, an average of approximately 36% of transactions involved two publicly traded corporations, 48% were public companies acquiring private companies, and 16% were private companies acquiring public companies. After 2010, the breakdown is 26%, 52%, and 22%, respectively.

2.3 Employing LLMs to Study M&A Motivations

This paper employs a text classification approach similar to [Hansen and Kazinnik \[2023\]](#). At a high level, the process for using an LLM to evaluate different transactions is the same as if one was asking a human research assistant to do the same task. A researcher provides written instructions, along with any documents that provide context, to the LLM and has it process the classification problem. Like a human research assistant, the LLM is not perfect and the researcher must check its work. The specific LLM employed is the GPT-4 model⁸ developed by OpenAI. During the model’s setup stage, the model is provided with a PDF copy of the 2023 U.S. merger investigation guidelines⁹ as well as the instructions

You are an expert working for the U.S. government evaluating merger and acquisition (M&A) transactions. Your job is to determine if a transaction’s parties are horizontally or vertically linked. A transaction does not need to

⁸Specifically, the gpt-4-0125-preview model

⁹[U.S. Department of Justice and the Federal Trade Commission \[2023\]](#)

be anticompetitive to have linkages. To support your analysis, strictly adhere to the information in the provided statement. Do not use outside general knowledge.

The input will be of the form:

Id number: The transaction's Id number

Company A: Company A's description.

Company B: Company B's description.

You may be provided multiple inputs. These are unrelated transactions.

Ouput:

Id number, horizontal linkage code, vertical linkage code

The horizontal linkage code values are:

1 there is no potential horizontal linkage

2 there is potential horizontal linkage

3 the information provided does not allow you to determine if there is a potential horizontal linkage

The vertical linkage code values are:

1 there is no potential vertical linkage

2 there is potential vertical linkage

3 the information provided does not allow you to determine if there is potential vertical linkage

Do not provide any additional details except the id number and linkage code.

Then, the GPT model is provided with a textual description of both parties in the M&A transaction. One limitation of this approach is that the GPT model is extremely sensitive to changes in the prompt. When the sentence "Do not use outside general knowledge" was removed, it would use information outside the text descriptions¹⁰. Also, if it is given an option of "uncertain" along with "yes" and "no", it will default to that when it should otherwise respond with "no". The reasoning it frequently gives is that it is unsure if the statement is a complete representation of the company. While this conservative approach may be appreciated, it is not useful for this application.

2.3.1 Example Transactions

To start, consider the 2004 acquisition of Where 2 Technologies by Google¹¹, the textual description of Where 2 Technologies is

Where 2 Technologies provides mapping-related startup technologies. It was co-founded in 2003 by Jens and Lars Eilstrup Rasmussen in Sydney.

and the textual description of Google is

¹⁰For example, with the Where 2 Technologies and Google transaction, the model would reply that Where 2 and Google compete since Google has Google Maps. This is incorrect since the Where 2 Technologies acquisition was the start of Google Maps. The model was relying on information not included in Google's textual description (this is detailed further in the next subsection).

¹¹SDC Master Deal Number 1949520040

Google Inc, located in Mountain View, California, provides Internet search engine services and advertising solutions through its internet site and intranet solutions via an enterprise search appliance. It has local offices in Mississippi, Georgia, Texas, Colorado, Illinois, Massachusetts, North Carolina, and New York, and international offices located in Belgium, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Ireland, Italy, the Netherlands, Norway, Poland, Portugal, the Russian Federation, Spain, Sweden, Switzerland, Ukraine, the UK, Canada, Latin America and the Middle East. The company has automated search technology that helps people to obtain access to relevant information from vast online index. Trademarks include Google, YouTube, DoubleClick, DART, AdSense, AdWords, Gmail, I'm Feeling Lucky, PageRank, Blogger, orkut, Picasa, SketchUp and Postini. It can be accessed through destination sites at Google.com and 112 other international domains, such as Google.ba, Google.dm, Google.nr, Google.co.jp and Google.ca. The Google interface is available in 120 languages. The company was founded in 1998. 120 languages. The company was founded in 1998.

The GPT model responds that it does not believe there is a horizontal linkage, but there is a vertical linkage. It's reasoning is

The determination that there is a vertical linkage between Google Inc. and

Where 2 Technologies is based on the nature of the services and products each company offers. Google Inc. provides a wide range of internet services, including search engine services, advertising solutions, and a variety of other internet-related technologies. Where 2 Technologies, on the other hand, specializes in mapping-related startup technologies.

The vertical linkage arises from the potential integration of Where 2 Technologies' mapping technology into Google's broader suite of internet services. Specifically, Google could utilize Where 2 Technologies' mapping technology to enhance its own products, such as improving its search engine's ability to provide location-based search results or integrating mapping technology into its advertising solutions. This kind of relationship, where one company's output (in this case, mapping technologies) is used as an input in another company's product or service (Google's internet services), is indicative of a vertical linkage according to the merger guidelines.

The reasoning follows the merger guidelines logically. Google, instead of simply forwarding their customers to different online map services, integrated mapping into their set of services and was able to capture more of their customers' digital interactions.

Next, consider the 2014 acquisition of WhatsApp Inc. by Facebook Inc. (now Meta

Platforms)¹². At the time, the description of Facebook Inc. was

Facebook Inc, located in Menlo Park, California, provides online social networking services. It primarily provides a Web-based platform that facilitates sharing of information, including media files, but which principally serves as a personal networking tool. It has over 800 million users, and has offices in several locations including Atlanta; Birmingham, Michigan; Chicago; Dallas; Detroit; New York; Venice Beach, California.; Washington, DC; Austin; Seattle; International offices: Dublin; Hamburg; Hong Kong; Hyderabad; London; Madrid; Milan; Paris; Selangor; Singapore; Stockholm; Sydney; Tokyo; and Toronto. The Company was founded in February 2004.

and the description of WhatsApp Inc. was

WhatsApp Inc, located in Santa Clara, California, develops cross-platform mobile messaging application software. It was founded in 2009.

The GPT model responds that the merger has horizontal linkages but not vertical linkages. When asked to explain its answer, the model stated

Facebook Inc. and WhatsApp Inc. are horizontally linked because both companies operate in the digital communication and social networking space, albeit with different focuses-Facebook on broad social networking and WhatsApp on mobile messaging. This positions them within the same industry,

¹²SDC Master Deal Number 2611055020

potentially competing for user engagement and time spent on digital communication platforms. There is no vertical linkage, as there's no indication that one company's services serve as inputs for the other's operations, which would be necessary to establish a supplier-purchaser relationship within a supply chain.

The model has again logically reasoned given the information it was presented. The model understood that social networking and messaging can be substitutes.

2.3.2 Evaluating the Model's Accuracy

A large language model, such as GPT-4, is a black box. There is no formula that provides a direct measure of its accuracy. Instead, a large language model's accuracy must be evaluated in the same manner that a researcher would evaluate the accuracy of a human research assistant: manually double-checking a subset of the labeled transactions. Of the transactions submitted for classification, a random sample of 200 was drawn. These are manually labeled and compared to the predicted value. Table 2.2, displays the confusion matrix which describes the relationship between the predicted labels and the test labels.

The confusion matrices are converted into two common performance metrics. The first is precision, the ability of the classifier to distinguish between the target class

and the other classes, defined by

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}.$$

The second is recall, the ability of the classifier to recognize members of the target class, defined by

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}.$$

Table 2.3, displays these accuracy statistics for both the vertically linked transactions and horizontally linked transactions. This reveals that the biggest weakness of this method is transactions that should have been labeled as uncertain. These transactions are primarily labeled as uncertain since there is not enough information from the textual description about one party’s business. For example, several buyers are private equity firms. From their textual descriptions, it’s unclear if they were previously involved in the acquired business sector. If restricted to transactions not labeled uncertain, this method is highly accurate. It just struggles on transactions with limited information. It is possible that a better prompt engineer could result in the more accurate tool.

2.4 Summary of Classification Results

. One downside of the GPT model is that it is expensive. As this study is an initial pilot, it is too costly to attempt the classification method on every transaction in the SDC Platinum data set at this time. Instead, one thousand transactions are sampled for every year¹³ resulting in a sample of thirty thousand transactions. The reason for sampling by year is to examine the change over time and check if the method is robust to different time periods. Once a final methodology is established, this can easily be applied to the entire data set.

The GPT model labels approximately 95% of the transactions in the sample¹⁴. Table 2.4 displays the results of the classification exercise by count shares and Table 2.5 displays the results by valuation share. Starting with counts shares, approximately 39% of transactions are horizontally linked, 39% of firms are vertically linked, and 6% of transactions have both horizontal and vertical linkages. Around 17% of firms had no vertical or horizontal linkages. While the potentially anti-competitive nature of horizontal mergers makes them the focus of M&A research, this result shows that vertical mergers are equally common. In terms of valuation, the shares of horizontally linked and vertically linked transactions are roughly the same as their count shares. There is an increase in the share of value where the parties were both horizontally

¹³The year the deal was announced.

¹⁴A transaction is unlabeled if the classifier returns uncertain for either horizontal or vertical linkages.

and vertically linked and a corresponding decrease in value for neither.

Focusing on different target firms' industries, the count shares are similar to the shares for the total sample. However, there are significant differences in terms of valuation shares. The value of transactions where target firms were in the tech sector is more concentrated in the vertically linked transactions. Pharmaceuticals and oil each have a larger share of value in horizontal mergers than their count share. Moving to whether the parties are public or private, when only the acquiring company is public, the count and value of shares are roughly similar to the total sample. However, if only the target company is public, these transactions are primarily vertical in nature. This suggests private companies are consolidating their supply lines. A similar result holds when neither company is public.

Since, in the United States, antitrust policy is primarily a function of the executive branch of government, it is possible that the composition of M&A transaction types may change depending on the political leanings of the administration. Figure 2.3, plot the transaction and value shares of M&A transactions by presidential administrations. Panel A shows that, by count shares, the composition of M&A transactions are relatively similar for all four administrations covered in the data set. Horizontal only mergers made up a slightly larger share — 40% vs an average of 35% in the other three administrations. Panel B shows that the value share of horizontal-only mergers increased for the Obama and Trump administrations: an average of 50%

vs. 33%. This increase in horizontal merges value share is caused by the value share of transactions involving both horizontal and vertical linkages decreasing from an average of 17% to 3% and transactions with only vertical linkages decreasing from an average of 40% to 35%.

2.5 Conclusion

This study's methodology is a novel approach to studying the relationships between parties in a M&A transaction. Not only does it push the frontier of machine learning in economics, it also provides database of M&A relationships that can be used in research on M&A transactions. There are multiple areas warranting future study. The GPT prompt is the key input to the classification exercise. Designing these prompts is more an art than a science. It is probable that a different prompt may improve the classification accuracy. However, there is no obvious way to achieve this goal. An additional source of further study is including additional labels in the classification exercise. The horizontal and vertical labels were the obvious starting point for M&A analysis. Additional characteristics, such as geographic overlap or common labor market, can be added with minimal effort. These would deepen our overall understanding of M&A activity. On the data front, while SDC platinum provides a nearly complete sample for M&A transactions involving publicly traded companies, it is less clear how representative it is for private companies. Better data

could provide better insights.

2.6 Tables

Table 2.1: M&A Transaction Summary Statistics

	All Transactions	Both Public	Public Buying Private	Both Private
Number of Transactions (Thousands)				
Total	182	43	71	41
Average Per Year	6	1	2	1
Total With Valuation	65	24	25	4
Transaction Value (Millions of 2017 Dollars)				
Mean	385	695	98	152
Standard Deviation	2616	3846	396	4401
p25	9	18	7	3
p50	35	75	22	8
p75	147	303	68	31
Target Firm's Industry (Percent of Total Valuation)				
Finance	19	20	20	34
Tech	14	13	22	6
Communications	13	14	5	8
Pharmaceuticals	8	9	6	1
Oil	6	6	5	5
Other Manufacturing	17	17	16	15
Other	22	21	25	31
Public Company (Percent of Transactions)				
Acquirer Only	40	0	100	0
Target Only	15	0	0	0
Both	24	100	0	0
Neither	22	0	0	100

Note: Industries are the target firm's self-selected Standard Industry Classified Code (SIC). Finance is any code beginning with 6. Tech is any code beginning with 73, 357, or 367. Communications' codes begin with 48. Pharmaceuticals' codes begin with 283 or 384. Oil's codes begin with 13. Other manufacturing is other codes beginning with 2 or 3. Firms included in this calculation must have a disclosed valuation.

Table 2.2: Classification Confusion Matrix
Horizontally Linked

		Predicted		
		True	False	Uncertain
Labeled	True	74	0	0
	False	5	78	0
	Uncertain	24	11	7

Vertically Linked

		Predicted		
		True	False	Uncertain
Labeled	True	60	14	0
	False	8	69	0
	Uncertain	18	26	3

Table 2.3: Classification Precision and Recall

Horizontally Linked

	Precision	Recall
True	0.72	1.00
False	0.88	0.94
Uncertain	1.00	0.17

Vertically Linked

	Precision	Recall
True	0.70	0.81
False	0.63	0.90
Uncertain	1.00	0.06

Table 2.4: Count Shares of M&A Transaction

	Horizontal Only	Vertical Only	Both	Neither
Total Sample	39	39	5	17
Target Firm's Industry				
Finance	46	34	5	15
Tech	35	41	6	18
Communications	51	34	4	12
Pharmaceuticals	38	36	10	16
Oil	46	37	6	11
Other Manufacturing	36	39	8	17
Other	37	40	6	16
Public Company				
Acquirer Only	39	38	7	15
Target Only	31	43	4	21
Both	40	40	8	13
Neither	41	34	4	20

Note: Industries are the target firm's self-selected Standard Industry Classified Code (SIC). Finance is any code beginning with 6. Tech is any code beginning with 73, 357, or 367. Communications' codes begin with 48. Pharmaceuticals' codes begin with 283 or 384. Oil's codes begin with 13. Other manufacturing is other codes beginning with 2 or 3. Firms included in this calculation must have a disclosed valuation.

Table 2.5: Value Shares of M&A Transaction

	Horizontal Only	Vertical Only	Both	Neither
Total Sample	38	38	12	10
Target Firm's Industry				
Finance	32	45	12	11
Tech	27	55	6	12
Communications	55	17	23	4
Pharmaceuticals	55	30	8	7
Oil	62	28	4	7
Other Manufacturing	35	31	17	17
Other	30	50	9	11
Public Company				
Acquirer Only	37	43	10	11
Target Only	22	58	1	19
Both	44	31	17	7
Neither	16	57	1	26

Note: Industries are the target firm's self-selected Standard Industry Classified Code (SIC). Finance is any code beginning with 6. Tech is any code beginning with 73, 357, or 367. Communications' codes begin with 48. Pharmaceuticals' codes begin with 283 or 384. Oil's codes begin with 13. Other manufacturing are other codes beginning with 2 or 3. Firms included in this calculation must have a disclosed valuation.

2.7 Figures

Figure 2.1: M&A Transaction Counts and Median Value By Year

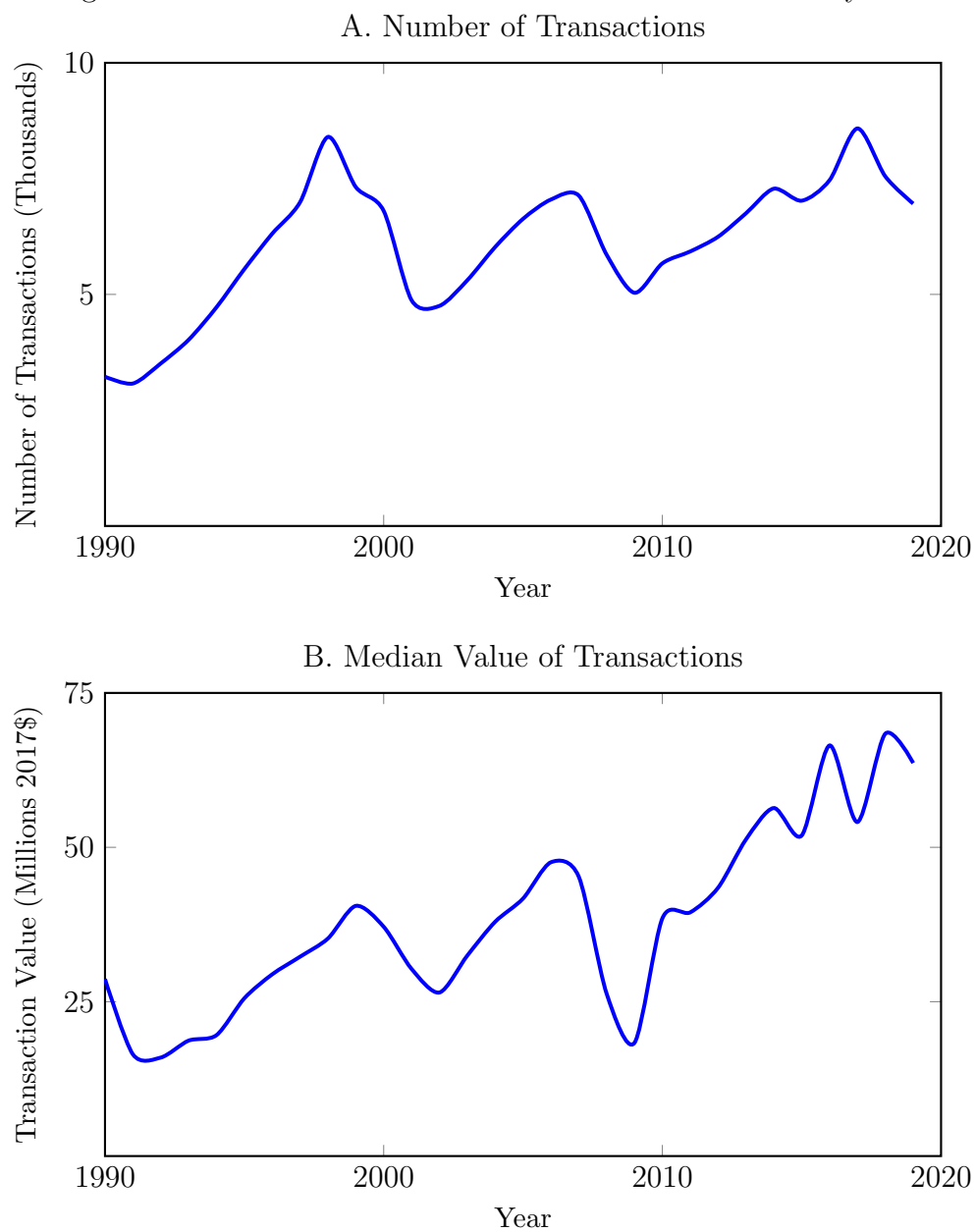
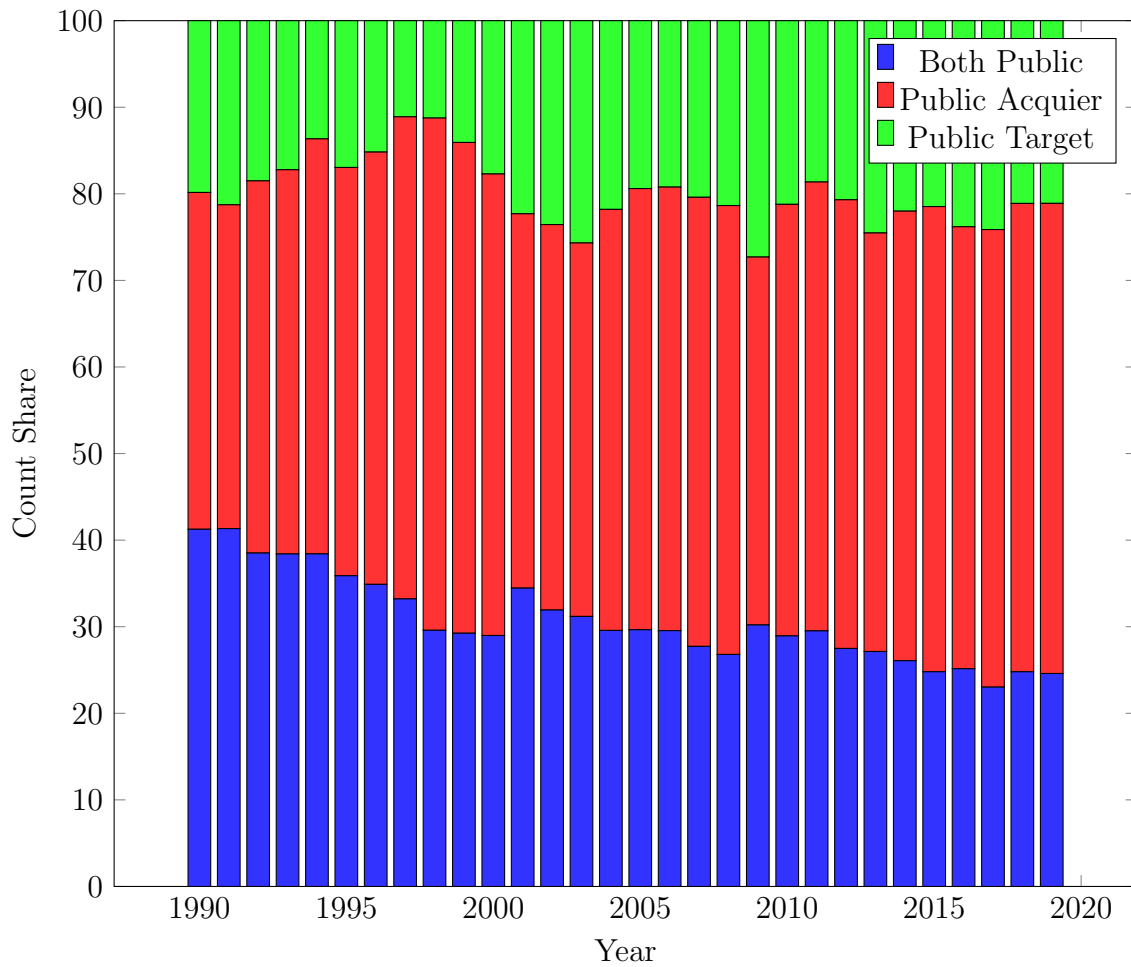
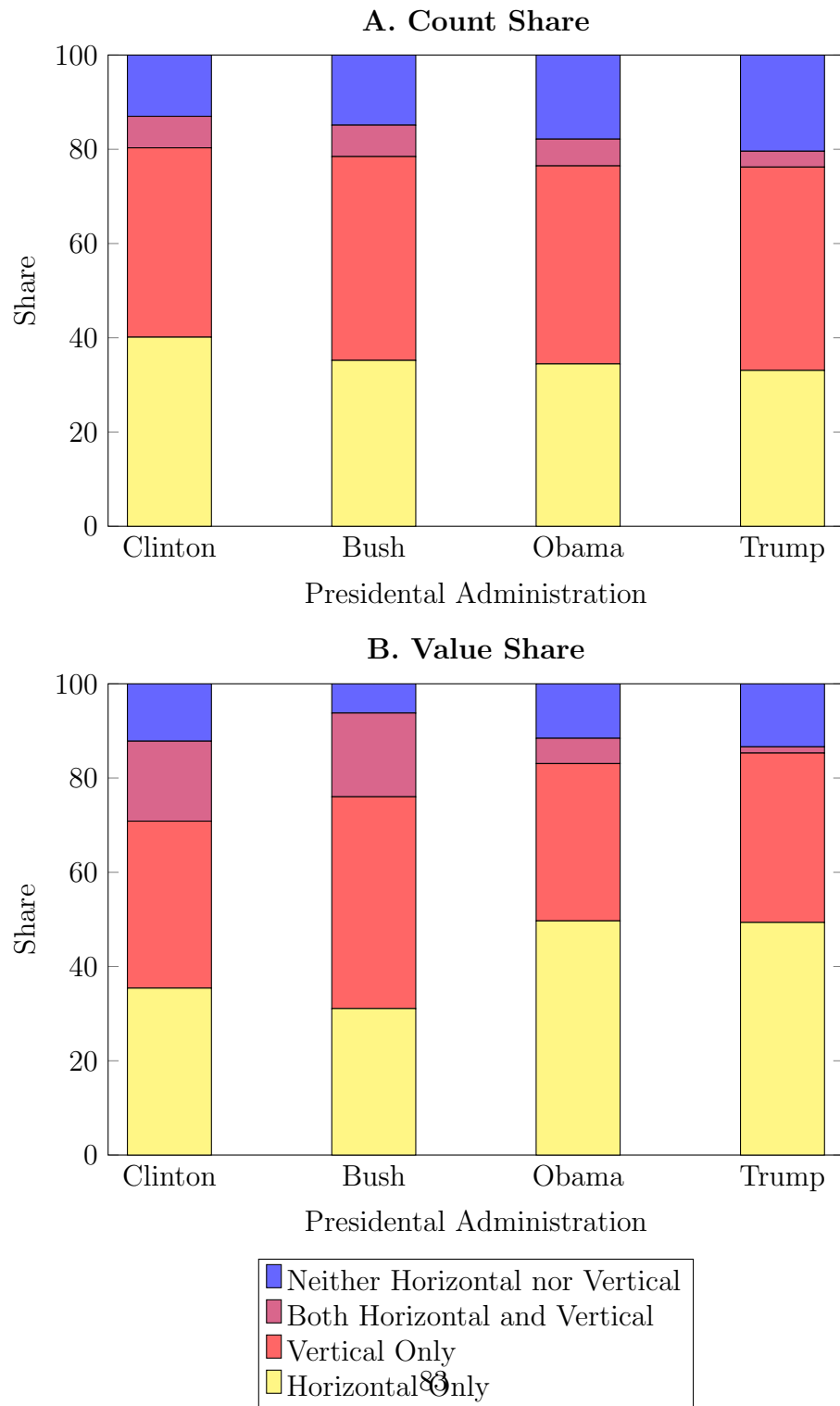


Figure 2.2: Shares of Transactions



Note: Both public means that both parties to the transaction are publicly traded. Public acquirer means only the acquiring company is publicly traded and the target company is private. Public target means the acquiring company is private and the target company is public.

Figure 2.3: Transaction Shares by Presidential Administration



Chapter 3

On the Nature of Entrepreneurship

Note: This chapter comes from joint work with Anmol Bhandari, Tobey Kass, Ellen R. McGrattan, Evan Schulz where May was a principle contributor. This project was facilitated through the Joint Statistical Research Program of the Statistics of Income Division of the United States Internal Revenue Service. May and McGrattan are IRS employees without pay under an agreement made possible by the Intragovernmental Personnel Act of 1970 (5 U.S.C. 3371-3376). This research was conducted while Kass was an employee at the U.S. Department of the Treasury. Any findings, interpretations, opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views or political positions of the Internal Revenue Service or the U.S. Department of the Treasury, or the National Science Foundation. All results have been reviewed to ensure that no confidential information is disclosed. All data work for this project involving confidential taxpayer information was done at IRS facilities, on IRS computers, by IRS employees, and at no time was confidential taxpayer data ever outside of the IRS computing environment.

3.1 Introduction

Despite volumes written on the topic, there is surprisingly little concordant evidence about returns to entrepreneurship.¹ The goal of this paper is to fill in part of the gap in our knowledge using U.S. administrative tax filings of employees and business owners over the period 2000–2015. We develop an econometric framework to estimate growth in incomes and use it to compare life-cycle income profiles and switching behavior of individuals who share similar characteristics but differ in their choice of self- or paid-employment. We then use these statistics to inform economic theories of occupational choice and entrepreneurship.

We construct life-cycle income profiles for groups of individuals with different demographic and labor-market characteristics. We utilize the Statistics of Income (SOI) Databank, which combines records from the Social Security Administration (SSA) and the Internal Revenue Service (IRS), providing us with demographic information such as age, gender, marital status, and the number of children as well as information on employment status, occupation, industry, own incomes, and household incomes. We use machine learning algorithms to impute additional information such as educational attainment and occupation to measures a broad notion of skills. Our measure of self-employment income is the sum of incomes from proprietorships (Form 1040, Schedule C net profits), partnerships (Form 1065, Schedule K-1 ordinary business

¹See [Parker \[2018\]](#) for a comprehensive review of the literature.

income), S corporations (Form 1120-S, Schedule K-1 ordinary business income), and own-business compensation (Form W-2 wages). Paid-employment income is wage income (Form W-2) less any own-business compensation. Along with the income measures, we use auxiliary data such as number of employees and other business deductions to classify individual-year pairs as either self-employed, paid-employed or non-employed.

Our first exercise is to estimate a rich specification of income over lifecycle that uses information on the group and cohort of individuals. It has three components. The first component is an individual-level fixed effect meant to capture latent abilities, preferences, and other unobservable characteristics. The second component is a time effect that depends on the individual's cohort and group and is meant to capture changes in income specific to our sample, such as the large recession occurring in 2008–2009. The third component is an age effect that depends on the individual's cohort and group and is meant to capture changes in income over the life cycle as individuals gain more experience on the job. Our identification scheme assumes age effects are similar across binned cohorts and, with differenced income from our dataset, allows us to estimate the time and age effects for all subgroups.

To overcome issues related to compositional bias, we separately study individuals who switch frequently between self-, paid-, and non-employment and those who are more "attached" in their employment status. By attached, we mean that the

individual has the same employment status for twelve or more years, with at most two switches in status over the 16-year sample period and no intermediate spells of non-employment. Contrasting income profiles of attached with those less that are less attached allows us to discern pecuniary versus non-pecuniary motives both of which could underly an individual's choice of being attached to an activity.

For the full sample of attached employees and entrepreneurs we find that growth in employee income declines across the life cycle, while growth in entrepreneurial income remains persistently high until mid-career and then gradually declines. At age 25, the mean incomes are \$34 thousand (in 2012 U.S. dollars) for the paid-employed and \$42 thousand for the self-employed. By age 55, the self-employed are earning more than twice the paid-employed, roughly \$210 thousand versus \$89 thousand. If we decompose the aggregate differences in income growth between the self- and paid-employed in their 30s and attribute them to subgroups of our sample, we find the key contributors driving these differences are married males with occupations requiring education and interpersonal skills that have jobs in health care, professional services, finance, retail trade, and construction. For these subgroups, the growth is highest during the mid-30s, suggesting that business owners make initial investments to build a business (as in [Bhandari and McGrattan \[2021\]](#)) or experiment early in their careers in order to learn their productive capabilities in different occupations (as in [Jovanovic \[1982\]](#)). Investment and experimentation would delay growth and generate

the delayed hump-shaped growth profiles that we observe. We also find that the profiles of individuals who have non-employment spells and are relatively attached to non-market work—a large fraction of our sample—are flatter and substantially lower than those attached to paid-employment or self-employment.

Our findings on life-cycle growth rates for self-employed individuals are different from those in the survey-based literature (See [Hamilton \[2000\]](#), [Hurst and Pugsley \[2011\]](#)) who find flatter profiles and conclude a large non-pecuniary role for self-employment. We reconcile those differences by comparing distributions of income by age and employment across the Current Population Survey with the IRS administrative data. We find that for comparable self-employed individuals, the average based on IRS data is as large as \$42 thousand more than the estimate based on CPS data. In contrast, the paid-employed averages are barely different across datasets for prime-age individuals. Furthermore, the CPS-IRS differences in median incomes by age and employment status are much smaller. We conclude that surveys fail to capture well the right-tail of income distribution, which for self-employed, contains most of the income. Thus, while the survey-based analysis paints a reliable picture of the median self-employed individual, it is not reflective of how the median dollar in self-employment is earned.

After studying income profiles, we analyze entry into and out of self-employment over the life cycle to gain a better understanding of entrepreneurial choice. For our

sample, exit rates decline significantly over the life cycle, with experimentation in entrepreneurship occurring at younger ages, but are flat across time. Entry rates are flat over both the life cycle and across time. Remarkably, we see little change during the 2008–2009 recession, which suggests that entrepreneurship was not used as a fall-back option. Relatedly, if we compare past labor incomes for observationally similar individuals—one entering self-employment and the other not—we find the newly self-employed had higher past income, which is inconsistent with the view that “misfits” are pushed into entrepreneurship. If we instead compare past asset incomes for these observationally similar individuals, we find the opposite: the newly self-employed had lower past asset income, which is inconsistent with the view that entrepreneurs face liquidity constraints. (See [Evans and Leighton \[1989\]](#).) If we compare earnings of those that switch employment status—whether they are switching from self- to paid-employment or vice versa—with those observationally similar peers that do not switch, we find nearly as many increases as decreases in income, suggesting that both pecuniary and non-pecuniary motives drive occupational choice.

After analyzing the mean growth rates by age and dynamics in and out of self-employment, we study the variability and persistence of income changes to investigate the risky nature of entrepreneurship. We focus on the dispersion and autocorrelation of income changes—two statistics that have been used in previous work to quantify the gains of greater insurance against idiosyncratic risk. If we measure dispersion

using the 90–10 difference in percentage growth (after netting out time and age effects), we find that the volatility is 3 times greater in self-employment relative to paid-employment, while the autocorrelations of the rates are roughly the same across employment status. While greater variability translates into higher welfare gains for smoothing entrepreneurial risk, we find that individuals in our sample have means to smooth consumption expenditures through spousal wages and other household income.

After documenting the key empirical patterns of our sample, we use theoretical predictions of an occupational choice model to interpret our findings. In the model, our theoretical entrepreneurs spend some time investing in self-created intangible assets—for example, customer bases and tradenames—and growing to an optimal size. There are risks in self-employment and young entrepreneurs start with little to no financial assets or other incomes that can be used to smooth consumption during the first years. Meanwhile, productive abilities must be learned and when they are, exit due to selection occurs. If exit does occur, the business is sold, intangible assets are transferred, and the owner switches to paid-employment.

Because we are interested in the role of investment and experimentation in generating realistic growth profiles and hazard rates corresponding to entrepreneurship, we compare model simulations to the youngest cohort of our sample—those born between 1970 and 1975—that are self-employed for at least five consecutive years prior

to age 35. For the simulations, we use the baseline parameterization of [Bhandari and McGrattan \[2021\]](#), who abstract from learning, and then use moments from the IRS subsample to set parameters of the learning process and income shocks. This parameterized version of the model is shown to generate profiles consistent with young entrepreneurs in our IRS sample. We find that learning is a necessary feature of the model: if there is too much certainty about business owners’ productive capabilities and the nature of business risk, then occupational choices are made quickly. In that case, the model cannot rationalize self-employment stints as long as five years followed by a switch. Similarly, we find that firm-specific investment is a necessary feature: if an owner only requires factor inputs that can be rented or hired without delay, then the business can be scaled to its optimal size immediately. In this case, the model cannot rationalize persistent differences in income growth when comparing profiles for entrepreneurs that continue in business and those that exit.

An important by-product of our work is a longitudinal database of business owners that can be used to develop predictive tools—both theoretical and statistical—for improved tax administration. This database allows for a broader scope of analysis, beyond what is possible with survey data alone. With surveys, researchers can study the typical entrepreneur, while we can study the typical dollar earned in self-employment and can track the individual earning it over a long period. What our analysis shows is that the typical dollar is earned by those with incomes in the top 25

percent that are attached to self-employment, and these individuals have life cycle income profiles that are easily distinguishable from their paid-employed peers.

3.2 Data

In this section, we describe our main sample drawn from U.S. administrative tax records.² We start with details of the data source and definitions of self- and paid-employment income. We then describe algorithms to impute skill and education levels.

3.2.1 Sample

When constructing our sample, we start with records in the SOI Databank, which is a de-identified balanced panel of all living individuals with a U.S. Social Security number over the period 1996 to 2015.³ For each individual there are rows, one for each year, and columns recording demographic information from the SSA (such as age and gender) and economic data from tax filings (such as information on individual income tax forms and attachments). This database is our primary source for data. The SOI Databank includes information on wages and salaries reported to the IRS on Form W-2 for employees and household-level Schedule C income reported on

²Replication codes and detailed documentation are available at the IRS.

³See [Chetty et al. \[2018\]](#) for full details on this database. We remove any person from our sample who died prior to 2015.

Form 1040. For individual proprietors, we assign incomes separately by Social Security number. For owners with pass-through businesses—partnerships and S corporations—we merge in information from Schedule K-1 filings attached to Form 1065 and 1120-S, respectively.⁴ The Schedule K-1 data is available since 2000 and thus our sample period ranges from 2000 to 2015. Because self-employment income must be reported on the standard Form 1040 when filing individual income taxes, we exclude from our baseline sample any individuals that exclusively use the simpler Forms 1040A or 1040EZ.

To construct income profiles by age, we use records for all individuals between the ages of 25 and 65 in the SOI Databank for the years 2000 through 2015—namely, birth cohorts 1950 through 1975. This balanced panel includes roughly 128 million individuals for 16 years (that is, 2 billion person-year observations).⁵ Another restriction we place on the sample is the availability of occupational information, which is used to impute levels of education and skill that play an important role in income determination. This restriction narrows our sample to roughly 80 million individuals. Details of the imputations are provided below in Section 3.2.3.

⁴Business net incomes of Subchapter C Corporate shareholders are not passed through to individual income tax forms until the companies distribute dividends or capital gains. Until then, there are no administrative records that can be attributed to the individual owners.

⁵The full SOI Databank sample over 2000-2015 has 7 billion observations, with 3 billion for ages 25 to 65 if we include individuals that are not in our sample all years.

3.2.2 Income Measures

For each individual-year observation, we compute two sources of income. The first is a measure of *self-employment income* and is defined as the sum of net profit or loss of sole proprietors (Form 1040, Schedule C, Line 31), the individual's share of ordinary business income from partnerships (Form 1065, Schedule K-1, Part III, Line 1), the individual's share of ordinary business income from S corporations (Form 1120-S, Schedule K-1, Part III, Line 1) and finally the individual's income paid by the S corporations that they own as wages (Form W-2, Box 1).⁶ The second is a measure of *paid-employment income* and defined as the wages and salaries (on Form W-2, Box 1) paid by businesses that are not owned by the wage-earner. We refer to the sum of self- and paid-employment income as *total income*, although it does not include other categories of adjusted gross income on the tax forms. These measures are computed before tax and transfers, exclude most employer fringe benefits, and are deflated by the Bureau of Economic Analysis's (BEA) personal consumption expenditure price index and reported in thousands of 2012 U.S. dollars. No adjustments are made to account for potential income underreporting.

Although individuals can have both paid-employment income and self-employment income, we assign individuals to distinct employment categories each year based

⁶Here, we omit capital gains as a source of self-employment income, although we acknowledge that there could be realized gains from sales of intangible assets reflecting entrepreneurial investment that should be included with self-employment income. Including such gains only strengthens our main finding.

on a test designed to gauge their primary activity. To do that we construct three categories: self-employed (SE), paid-employed (PE), or non-employed (NE) using the following definitions. An individual-year pair is classified as *self-employed* (SE) if the absolute value of self-employment income exceeds \$5 thousand (in 2012 dollars) and any of the following condition is true: (i) the absolute value of their self-employment income is greater than their paid-employment income; (ii) the sum across businesses of the individual’s ownership share times the number of its employees is larger than 1; or (iii) the sum across businesses of the individual’s ownership share of gross profits (receipts less cost of goods sold) are in excess of the individual’s paid-employment income. We take the absolute value of the income because young entrepreneurs incur significant expenses when building up their businesses and many have losses. The second additional criteria is added because hiring employees is indicative of owner attachment to self-employment. The third criteria allows for the fact that many successful business owners pay themselves little to minimize taxes but earn incomes later when selling their businesses.

Our notion of self-employment is distinct from papers such as [Smith et al. \[2019\]](#), [DeBacker et al. \[2022\]](#), [Garin et al. \[2022\]](#), and [Lim et al. \[2019\]](#), who all use IRS data to study business incomes. [Smith et al. \[2019\]](#) classify all individual recipients of K-1 as self-employed. Our definition excludes 43 million of the 138 million individual-year K-1 recipients in our sample from being classified as self-employed. These are

cases in which an individual probably spends very few hours running a business and receives very little income from business filings. While this is not a concern for top incomes, which is the focus of [Smith et al. \[2019\]](#), our focus is to learn about returns to entrepreneurship. Therefore, we deliberately use a more conservative test when categorizing entrepreneurial activity. [DeBacker et al. \[2022\]](#) use a panel that tracks tax filers for up to 32 years using the SOI sample from 1987. While this has the benefit of being a long panel, the number of self-employed individuals that are studied shrinks down to about 2,000 observations over a few cohorts. Such a restrictive sample would be unsuitable for achieving our two main goals: (i) to calculate life cycle income profiles using overlapping cohorts to infer time and age effects and (ii) to understand the determinants of self-employment by comparing outcomes for narrowly-defined groups—some of whom enter self-employment and some of whom do not. [Garin et al. \[2022\]](#) focus on Schedule SE filers. This is not suitable for our analysis because it misses entrepreneurs who make losses and S corporation owners that do not file Schedule SE. This is a significant fraction of business owners. [XXX add the new stuff on gig workers here]

Next we define paid- and non-employed categories. An individual-year pair is categorized as *paid-employed* (PE) if it is not already categorized as self-employed, and if paid-employment income of the individual in that year exceeds \$5 thousand (in 2012 dollars).

An individual-year pair is categorized as *non-employed* (NE) if it is not already categorized as SE or PE.

To distinguish observations that are non-employed from those that are actually paid- or self-employed but missing in the SOI Databank, we use auxiliary data from Form 1040 and the individual's Form W-2, Schedule C, or Schedule K-1 if any of these filings are available. Consider wage earners first. We compute a *wage gap* as the difference in wages and salaries reported on Form 1040 and the aggregated Box 1 wages on Form W-2 for the individual and spouse. If this gap is less than \$1,000, we use the Form W-2 data for the individual. If the wage gap is greater than \$1,000 and the individual is not married filing jointly, then we use the Form 1040 wage. If the wage gap is greater than \$1,000 and the individual is married filing jointly, then we need to consider two cases. First, there may be Form W-2 filings for the spouse but not the individual. In this case, we use the Form 1040 wage less the spouse's total W-2 wages. Second, the Form W-2 filings may be missing for both the individual and spouse. In this case, we take pro-rata shares for the two spouses based on tax filings in other years showing a wage gap less than \$1,000. If such information is unavailable, we split the Form 1040 wages evenly for the two spouses. In the case of business owners, we follow the same procedures as for wage earners, but in this case we use the owner's Schedule C or Schedule K-1.

In the first column of Table 3.1, we provide summary statistics for our main sample. The full sample has 80 million individuals. Summing across all individual-year observations, we find the shares of self-, paid-, and non-employed are 8 percent, 74 percent, and 18 percent, respectively. They earn an average of \$49 thousand (in 2012 dollars) in combined paid- plus self-employment income, with a range across the distribution from \$6 thousand at the 10th percentile to \$92 thousand at the 90th percentile—roughly a factor of 15. If we sum up self- and paid-employment incomes across all individual-year observations in our sample, we find the share of self-employment income is equal to 13 percent.⁷ The mean paid-employed income lies between 50th and 75th percentiles of the distribution of PE income, while the mean self-employed income is closer to the 90th percentile of SE income, indicating a substantial right-skewness of the self-employment income distribution for the whole population.

3.2.3 Imputations for Skills and Education

A large empirical labor literature focuses on skills and education as determinants of income. In this section, we impute information of education and occupation which is not readily available in the tax data. We later use these estimates when analyzing subgroups of the population of tax filers.

⁷If we include all earnings when computing this share, we find that self-employed individuals earn 16 percent of the total income.

Skills

After signing and dating the tax form, individual tax filers and their spouses are asked to self-report their occupation, which is summarized in the IRS data as a character string. The occupation information is available for e-filed returns for tax year 2005 and later, with the exception of 2012. For the sample of individuals born between 1950 and 1975, there are 89 million individuals that e-filed at least once in the years that these occupation strings are available. We are able to assign skill values to the subset of 80 million individuals in our main sample.

First, there are 73 million individuals that provide usable occupations, which can be mapped directly to a standard occupational classification (SOC) code.⁸ For these individuals, we assign skill values using the procedure of [Lise and Postel-Vinay \[2020\]](#). The idea is to create a mapping between the SOC codes assigned to individuals and their cognitive, interpersonal, and manual abilities. This is done with the aid of the Occupational Information Network (O*NET) summary of skill requirements needed for each occupation. Since the summary of requirements is long for each occupation, [Lise and Postel-Vinay \[2020\]](#) use a principal component analysis to construct indices—keeping the top three (orthonormal) components and ensuring that occupations requiring mathematics are encoded as “cognitive” skills, occupations requiring social perceptiveness are encoded as “interpersonal” skills, and occupations requiring

⁸We thank Raj Chetty and his team for providing us with a mapping between the strings and the SOC codes.

mechanical knowledge are encoded as “manual” skills.

Second, there are 7 million individuals for whom we impute a skill value.⁹ For these individuals, we apply a k -nearest neighbor classifier for the imputation using information on k “neighbors” from the subsample of the 73 million individuals that have a valid SOC code and assigned skill values. The neighbors share the same gender, marital status, birth cohort, 2-digit NAICS industry code and are nearest in paths of employment status and incomes.¹⁰ For each subgroup, we operationalize choosing near neighbors in the case of time-varying income variables by applying a principal component analysis (PCA) that maps a high-dimensional vector of statistics from our data to a lower-dimensional vector of moments. Inputs to the PCA are paid- and self-employment income in each sample year and moments of total income averaged across sample years. The specific moments are the mean, the standard deviation, the minimum, and the maximum, with the latter three normalized by dividing by the mean. The number of principal components depends on our choice of the fraction of variance to be explained, which we denote here by v . Thus, we have two parameters to choose: the number of neighbors k and the fraction of variance v —and we assume they are fixed for all subgroups.

We choose parameters to maximize the predictive accuracy of the k -nearest neighbor

⁹For instance, business owners might fill in “self-employed,” which is not a valid SOC code.

¹⁰In Section 3.3.3, we group individuals into six different categories of employment status based on attachment and type of work.

classifier.¹¹ To do this, we pull a random sample of subgroups and split them into three subsamples: 70 percent for training, 20 percent for tuning, and 10 percent for validation. For each (k, v) pair, we use the training data to train the classifier and make predictions for the tuning set. We use the validation data to test prediction out of sample. The result of this exercise is $k = 11$ and $v = 75$ percent. With these parameters, we apply the classifier to impute skills values for 7 million individuals without usable SOC codes.

Education

The only indicators of education in the IRS microdata are occupation strings with “student” and tuition payment statements (Form 1098-T) filed by eligible educational institutions starting in 1998. To ensure fuller coverage of college attainment, we use a classification algorithm and source data from the Annual Social and Economic Supplement (ASEC) of the Current Population Survey (CPS) to predict the likelihood of college attainment.

We define an individual as being “college-educated” if they have completed at least an associate’s degree—which would thus include bachelor’s, master’s, professional school, and doctorate degrees. All others are considered “not-college-educated.” For

¹¹More precisely, we use an F_1 score.

each year t , we run the regression:

$$\Pr(E_{it} = 1|X_{it}) = \text{CDF}(\beta_t X_{it}), \quad (3.1)$$

where $E_{it} = 1$ if the individual is college-educated and 0 otherwise for t between 1995 and 2020. The function CDF in (3.1) is the cumulative distribution function of the standard normal and variables included in X_{it} are as follows: gender; annual pre-tax wages and salaries; positive business income (equal to 0 if income is negative); negative business income (equal to 0 if income is positive); marital status; number of children (with separate variables for none, 1 child, and so on, up to 9 or more); 5-year birth cohort; SOC minor occupation code; and 2-digit NAICS industry code.¹² When we used 80 percent of our CPS sample each year to train the classifier and 20 percent to validate the predictions, we were able to correctly predict the education level with 75 to 80 percent accuracy. Coefficients from the CPS-trained classifiers then are used with microdata from the IRS to impute an education indicator for all tax filers in our sample of 25-to 65-year olds.¹³

In Table 3.1, we report that the fraction of individuals categorized as college-educated for the main sample in the first column under “Education and skills.” The imputation

¹²Some IRS tax filers do not have a valid NAICS code and do not have a SOC minor code. Additional regressions were run using (i) the SOC minor codes with no NAICS; (ii) the SOC major code and NAICS; and (iii) NAICS but no SOC.

¹³All variables in X_{it} are available in the IRS data, although the IRS occupation field is only available for tax years after 2005 and later (not including 2012) and is a string rather than a SOC code.

results indicate that 53 percent of the main sample is classified as college-educated, which is close to our estimate of the fraction that are cognitively-skilled. More individuals are categorized as interpersonally skilled (59 percent) than manually skilled (38 percent).

3.3 Measuring Returns to Employment

Our main goal is to understand the sources of differences in returns to self- and paid-employment and the implications for theory. To investigate the differences, we start by measuring how income from an activity—either self-employment or paid-employment—varies with age. We first outline the challenges to accurately measure the age profile of income. Then we describe how to use the novel features of our data with an econometric method that is designed to overcome these challenges.

3.3.1 Some Challenges

A natural starting point for measuring returns to employment is to specify a Mincer-type earnings regression and estimate average income by age after controlling for observables. This regression procedure is widely used and can be implemented with repeated cross-sections. (See, for example, [Hamilton \[2000\]](#).) There are several concerns with this approach. First, differences in average incomes could be driven by

selection—we are simply comparing individuals that differ in their latent characteristics. Second, differences in average incomes over age could be driven by the changing composition of the underlying groups.

We develop an econometric approach that addresses these two issues. First, we estimate income by age across activities allowing for an intercept whose distribution by individual characteristics—whether they are latent or observed—as well as calendar time is essentially unrestricted. Second, we use the long panel aspect of our data to classify individuals based on their attachment to an activity to mitigate problems with composition. We then separately study income-age profiles of individuals who are attached to their employment status by narrow skills/industry/demographic categories and income-age profiles of individuals who are less attached and transit into and out of self-employment.

3.3.2 Econometric Framework

We next describe and motivate the statistical model and estimation procedure that we use to estimate growth in incomes over the life cycle. Our method exploits the presence of multiple cohorts to separately estimate age and time effects for disaggregated subgroups within employment status. For now, we describe the procedure for an arbitrary assignment of individuals to groups and later describe how we construct the groups to minimize selection and composition bias.

We start with some notation. Let $i \in I$ be a set of individuals; $t \in \mathcal{T} = \{t_0, t_0 + 1, \dots, t_0 + T\}$ be a set of calendar dates; $c \in \mathcal{C} = \{c_0, c_0 + 1, \dots, c_0 + C\}$ be a set of birth years (or *cohorts*); $a \in \mathcal{A} = \{a_0, a_0 + 1, \dots, a_0 + A\}$ be a set of ages; and $g \in \mathcal{G}$ be a set of observable time-invariant characteristics (or *groups*) that partition I . Let $y_{i,t}$ be the income of individual i at date t . With slight abuse of notation, we use $a(i, t)$ to denote the age of individual i at date t , $g(i)$ to denote the group of individual i , and $c(i)$ to denote the cohort of individual i .

We define two functions $\beta : \mathcal{G} \times \mathcal{T} \rightarrow \mathcal{R}$ and $\gamma : \mathcal{A} \times \mathcal{G} \times \mathcal{C} \rightarrow \mathcal{R}$ that capture time, age, and cohort effects. We use the notation $\beta_{g,t}$ and $\gamma_{c,g}^a$ to denote the values of these functions for a particular collection of $\{g, t, a, c\}$, and $\beta_{g(i)}$ and $\gamma_{c(i),g(i)}^{a(i,t)}$ to be the values associated with an individual-time pair (i, t) . Consider the following specification for income

$$y_{i,t} = \alpha_i + \beta_{g(i),t} + \sum_{a=a_0}^{a=a(i,t)} \gamma_{c(i),g(i)}^a + \epsilon_{i,t}, \quad (3.2)$$

where $\epsilon_{i,t}$ is a disturbance term for individual i at date t . The model for income in equation (3.2) is quite rich. It has three components. First, the parameters $\{\alpha_i\}$ are the unobservable individual-level fixed effects that capture permanent aspects of latent ability, family inputs, and preferences as well as level effects tied to birth cohorts. We impose no restrictions on how these characteristics are distributed in the population or correlated with observable groups. Second, the parameters $\{\beta_{g,t}\}$

are the time effects that vary by calendar time and differ across groups. These parameters capture effects on income such as business cycle fluctuations. Third, the parameters $\{\gamma_{c,g}^a\}$ are the age effects that vary by age, cohort, and group. We are particularly interested in variations across subgroups based on employment status and other characteristics such as skills, industry, and demographics.

It is well-known and easy to see that one cannot separately identify β and γ from data on income. For instance, for a fixed group g , adding a constant to all $\gamma_{c,g}^a$ for which $c + a = t$ is observably indistinguishable from adding the same constant to all $\beta_{g,t}$. To make progress, we follow impose the following condition.

Condition 3. *Age-effects are the same across cohort bins of size $N_c \geq 2$.*

Below, we use the notation $\bar{\gamma}_g^a$ to indicate the age effect of a group g , which are now modified to include a specification of the cohort bin, say, individuals born in the 1950s, 1960s, or 1970s. It is worth pointing out that while we impose age effects for sets of cohorts to be the same, we impose no restrictions on how cohorts affect the level of income. The differences in mean income by cohort are absorbed in the fixed effect for individual i , namely, α_i . Condition 3 allows us to exploit the overlapping structure of our data to separate out age effects from time effects.

Next, we derive the formulas needed to implement the estimation procedure. Let Δ be the time difference operator so that $\Delta x_t = x_t - x_{t-1}$. Apply Δ to equation (3.2)

to obtain

$$\Delta y_{i,t} = \Delta \beta_{g(i),t} + \bar{\gamma}_{g(i)}^{a(i,t)} + \Delta \epsilon_{i,t}.$$

We work with differences in levels rather than in logarithms given that some businesses make losses and owners' income y_{it} can be negative.¹⁴ To estimate the age and time effects, we propose the following least squares problem:

$$\min_{\{\Delta \beta_{g,t}, \bar{\gamma}_g^a\}} \sum_{g \in \mathcal{G}} \sum_{t \in \mathcal{T}} \sum_{i \in \mathcal{I}} \left(\Delta y_{i,t} - \Delta \beta_{g(i),t} - \bar{\gamma}_{g(i)}^{a(i,t)} \right)^2. \quad (3.3)$$

By examining the first-order conditions of this minimization problem, we can better understand how the estimator works. Let $N_{g,t}^a$ be the number of individuals of group g , age a , at calendar date t . Let

$$\begin{aligned} \overline{\Delta y}_{g,t} &= \frac{\sum_{i \in \mathcal{I}: g(i)=g} \Delta y_{i,t}}{\sum_{a \in \mathcal{A}} N_{g,t}^a} \\ \overline{\Delta y}_g^a &= \frac{\sum_{t \in \mathcal{T}} \sum_{i \in \mathcal{I}: a(i,t)=a, g(i)=g} \Delta y_{i,t}}{\sum_{t \in \mathcal{T}} N_{g,t}^a} \end{aligned}$$

be the average income growth for group g between dates $t-1$ and t and the income growth averaged across time for individuals in group g between ages $a-1$ and a ,

¹⁴Later we discuss why our econometric procedure avoids issues arising from heteroskedastic errors.

respectively. We can rearrange the optimality conditions to get

$$\bar{\gamma}_g^a = \bar{\Delta y}_g^a - \sum_{t \in \mathcal{T}} \left(\frac{N_{g,t}^a}{\sum_{j \in \mathcal{T}} N_{g,j}^a} \right) \underbrace{\left\{ \bar{\Delta y}_{g,t} - \sum_{k \in \mathcal{A}} \left(\frac{N_{g,t}^k}{\sum_{\ell \in \mathcal{A}} N_{g,t}^\ell} \right) \bar{\gamma}_g^k \right\}}_{\Delta \beta_{g,t}}. \quad (3.4)$$

Equation (3.4) expresses $\{\bar{\gamma}_g^a\}$ as linear combinations of two summary statistics of data, $\{\bar{\Delta y}_g^a\}$ and $\{\bar{\Delta y}_{g,t}\}$ with weights $\{N_{g,t}^a\}$. Specifically, the age effects for some age a are given by the average income growth $\bar{\Delta y}_g^a$ for that age minus an appropriate weighted-average of the time effects $\{\Delta \beta_{g,t}\}$. The weights that appear in the adjustment correct for the possibility that the age distribution could be changing over time, which is relevant in our sample period.

To understand the intuition for the adjustment term in (3.4), consider the case in which the age distribution is constant across time, that is,

$$\frac{N_{g,t}^a}{\sum_{a \in \mathcal{A}} N_{g,t}^a} = \frac{\bar{N}_g^a}{\sum_{a \in \mathcal{A}} \bar{N}_g^a}, \quad (3.5)$$

where $\bar{N}_g^a = \sum_{t \in \mathcal{T}} N_{g,t}^a$. With some algebra, we can show that $\bar{\gamma}_g^a = \bar{\Delta y}_g^a - \bar{\Delta \beta}_g$, where $\bar{\Delta \beta}_g = \sum_{t \in \mathcal{T}} \Delta \beta_{g,t} / T$ is the average of time effects for group g . It simply says that the estimate of the age effect equals the average income growth for that age minus a simple average of the time effects. However, equation (3.5) does not hold in typical panel datasets and, therefore, the second term on the right-hand-side of

equation (3.4) gives the appropriate adjustment.¹⁵

We make two more observations about equation (3.4). First, the age effect $\bar{\gamma}_g^a$ can be estimated separately for each group g . Second, one can show that the rank of the system formed by stacking equation (3.4) for each age is $A - 1$. Therefore, we need an additional restriction—one for each group—to solve for the age effects $\{\bar{\gamma}_g^a\}$ uniquely. Following Hall [1968] and Deaton [1997], we impose the following condition.

Condition 4. *The average time effect satisfies*

$$\frac{\overline{\Delta\beta_g}}{\bar{y}_{g,t_0}} = \frac{\mu_g}{T} \sum_t (1 + \mu_g)^t \quad (3.6)$$

for some pre-determined constant μ_g , where $\bar{y}_{g,t_0} = \sum_{i \in \mathcal{I}: g(i)=g} y_{i,t_0} / \sum_{a \in \mathcal{A}} N_{g,t_0}^a$ is the average income for group g at the beginning of the sample.

Condition 4 allows the estimation to match the cyclical variation in the time effect across groups in a flexible way. This is especially helpful in our sample given the severe economic downturn in 2008–2009. In particular, we do not need to take a stand on the differential effects of aggregate shocks on groups.

¹⁵In our sample, we have a balanced panel and, therefore, the mean age is necessarily increasing in calendar time as the population is aging.

3.3.3 Groups

To implement the approach sketched out in the previous section, we need to define groups. A *group* is a Cartesian product of time-invariant characteristics that we call *subgroups*. In our case, there are 46,080 subgroups. In this section, we provide a summary of the subgrouping.

Given our interest in the returns to entrepreneurship, the two most relevant characteristics are: (i) how attached individuals are to market work, whether it is paid- or self-employment, and (ii) how attached a working individual is to self-employment. In Section 3.2.2, we assigned an employment status to each individual-year observation: “SE” for self-employed, “PE” for paid-employed, and “NE” for non-employed. To address compositional bias, we analyze income profiles by separately studying working individuals who change status and those who do not. We implement that by using the status variable across time to group individuals according to how attached they are to self- or paid-employment. An individual is labeled *attached* if we observe the same employment status for twelve or more tax years with two or fewer changes in employment status during the sample. To be included in the subsamples of attached self- or paid-employed, we also require that any switching in and out of self- or paid-employment not include intermediate years of non-employment. Those with years of non-employment are categorized separately below.

In Table 3.1, we report counts and characteristics in the second and third columns

for the subsamples of individuals that are attached to paid- and self-employment, respectively. There are 42 million in paid-employment, which is about 52 percent of the individuals. They earn 66 percent of total income, 76 percent of all paid income, and roughly 4 percent of entrepreneurial income. The attached self-employed, numbering 2 million individuals, comprise only 2 percent of the individuals in our main sample, but 8 percent of total income and 51 percent of entrepreneurial income.

While small in number, the entrepreneurs in attached self-employment have relatively high incomes when compared to sample totals. Their total combined income from self- plus paid-employment averages \$152 thousand—far more than the average attached paid-employed earning \$62 thousand. Given our definition of “attached,” it is reassuring to find that most of the total income for self-employed is indeed self-employment income and similarly so for the paid-employed with respect to paid-employment income. Comparing distributions of incomes for these two groups, we find more skewness in self-employment income, as expected. Later, we investigate this further when analyzing the longitudinal data across the life-cycle.

Separate results are reported for groups that do more switching in employment status. The *almost attached* groups have the same employment status for twelve or more tax years but switch more than twice between self- and paid-employment. The shares of the almost attached are much smaller and are analyzed separately as a robustness

check, but summary statistics in Table 3.1 show that they are similar in characteristics to their attached peers. The *mostly switchers* have twelve or more years in either self- or paid-employment—without an intermediate spell of non-employment—and experience at least five or more years in both types of employment. This group is similar in size to the attached self-employed, but those at the top of the distribution earn much less. This group will also be used to gain insight into motivation for entering and exiting self-employment.

The last category is *any non-employment* which includes individuals that have switched in and out of non-employment from self- or paid-employment at least once or individuals that have five years of non-employment during the sample period. This group is large in counts—roughly 41 percent of the total sample—and as a group account for a significant fraction of self-employment income, although they earn on average only \$21 thousand in both self- plus paid-employment.

If we were to compare the earnings to those of a full-time worker earning the average federal minimum wage for the year, converted to 2012 dollars, we would find that 57 percent have lower total incomes.

Even if we condition on those with only one year of non-employment, we find below-average total incomes. These individuals account for 15 percent of the any-non-employed individuals and earn \$43 thousand on average—with \$35 thousand from

paid-employment and \$8 thousand from self-employment. Furthermore, if we condition on the any-non-employed that are more attached to employment, we find that they are few in number and make less on average than those categorized elsewhere. For example, only 28 percent of the any non-employed do 12 or more years of paid-employment. Only 2.8 percent of the any non-employed do 12 or more years of self-employment. Those that do self-employment for at least twelve years have on average a total income of \$49 thousand in our sample—far less than groups with no intermediate spells of non-employment.

In addition to employment attachment, we use other observables to group individuals. The subgroup *College-educated* has two values: 1 if the education classifier was above the 0.5 cutoff and 0 if not. The subgroups *Cognitive*, *Interpersonal*, and *Manual* each take on one of two values: 1 if the skill value is above the 0.5 cutoff and 0 if not. *Industry* is the 2-digit NAICS code for the company paying the highest W-2 wages for the paid-employed individual or the company with the highest gross profits owned by a self-employed individual. The industry code takes on 21 possible values (including “missing”). When grouping individuals, we assign them the code observed in most sample years. The subgroup *Gender* has two values: “M” for male and “F” for female. The subgroup *Married* has two values: 1 if the individual is married for nine or more years in the sample—not necessarily to the same person—and 0 otherwise. The subgroup *Children* has two values: 1 if the individual has children

and 0 otherwise. The subgroup *Cohort* has 3 values: “1950s” if born between 1950 and 1959; “1960s” if born between 1960 and 1969; and “1970s” if born between 1970 and 1975. Since we are working with a balanced panel, we observe a significant overlap of cohorts over time, namely, 26 cohorts (birth years 1950–1975), across 41 ages (25–65) and 16 calendar years (2000–2015).¹⁶

Summary statistics for education, skills, industries, demographics, and other incomes are included in Table 3.1 for subgroups with different employment attachment. Relative to the sample total, all but the group with non-employment spells are more likely to be college-educated and have cognitive and interpersonal skills. There are notable differences in industrial composition across groups. Those with more self-employment attachment are found primarily in construction, professional services, health care, and other services, while the largest sector overall is manufacturing. The demographic data show that 50 percent of the sample is male. Most are married for a majority of years they are in the sample—about 64 percent—and most have children—about 82 percent. The median birth year for our sample is 1963. Across subgroups, we find uniformity in numbers of children and birth year, but some differences in gender shares and marital status.

The two largest groups—attached paid-employed and any-non-employed—have larger shares of women and the mostly unmarried. The any-non-employed were categorized

¹⁶In robustness exercises, we expand \mathcal{G} to include average income deciles in order to check if there are any issues arising from heteroskedastic errors in the baseline regression. We find that our main quantitative results are not sensitive to this change.

as a separate group in large part because of the many ways they differ from other subpopulations: on average, these individuals earn significantly less, are primarily low-skilled, and have a higher concentration of women than any other category. Not surprisingly, we find average household incomes—whether from spousal wages or asset incomes or transfers like unemployment insurance—are higher for the any-non-employed than for the employed groups in the sample.¹⁷

3.4 Results

In this section, we report on the estimated entrepreneurial income and growth profiles and then investigate factors impacting occupational choice. Our focus here is primarily on the attached self-employed, but central to our analysis are comparisons to the attached paid-employed with the same demographics, skills, and industries. These comparisons are relevant because the self-employed are rewarded for making firm-specific investments and would thus have potentially different income growth profiles. We also compare growth profiles of the attached self-employed to those for the mostly switching and any-non-employed, two groups that earn a significant fraction of self-employment income in the aggregate. Central to our analysis of occupational choice are comparisons of past wage, asset, and other household incomes for

¹⁷Asset income is the sum of the following categories reported on Form 1040: taxable interest, tax-exempt interest on municipal bonds, dividends, Schedule D profit or loss, and Schedule E profit or loss (other than distributions from S corporations and partnerships).

those that enter self-employment and those that do not. We investigate these choices for the full sample and for the youngest cohorts that experiment most. Finally, we relate our findings to the existing empirical literature and emphasize results most relevant for distinguishing between theories of occupational choice.

3.4.1 Entrepreneurial Incomes

We start with the main findings of the least squares estimation of equation (3.3). Estimates of the time effects provide a summary of cyclical growth patterns, including impacts during the 2008–2009 downturn. Estimates of the age effects provide a summary of life-cycle growth patterns for the self-employed, which can be compared to more familiar patterns of the paid-employed. Netting out time and age effects, we then analyze the residual growth, which contributes most to the variability of incomes and is thus relevant to understanding the nature of risk-taking in entrepreneurship.

Cyclical Growth

In Figure 3.1, we plot the time effects relative to average income for individuals that are attached, either to self- or to paid-employment, that is, a weighted sum of $\Delta\beta_{g,t}$ divided by $\bar{y}_{g,t}$ for each subgroup $g \in \mathcal{G}$ where \mathcal{G} is either the attached self- or paid-employed group and weights are constructed with subgroup counts.¹⁸ The values are

¹⁸In all results reported, we exclude the top and bottom 0.01 percent outliers. Our main aggregated results are not affected if we include them.

reported in percents and displayed for tax years 2001 to 2015. As expected, there is a decline in growth during the Great Recession, with paid-employed growth rates falling to a low of -1 percent and self-employed growth rates falling to a low of -16 percent. As we see from the figure, the flexibility of the econometric specification allows for differences in timing and magnitudes. Declines start earlier for the self-employed and are much larger in the midst of the downturn than those for the paid-employed. Interestingly, both groups see improvements by 2010, continuing on until 2012.

Life-cycle Growth

In Figure 3.2, we display the *integrated incomes* and associated growth profiles for the same sample of attached self- and paid-employed used in Figure 3.1. Panel A shows the weighted average integrated incomes. For each age $a \geq 25$, we compute average income for 25 year olds in group g plus $\sum_{j=26}^a (\bar{\gamma}_g^j + \overline{\Delta\beta}_g)$ and denote this sum for group g at age a by $Y_g(a)$. In Figure 3.2, we plot the averages using sample counts for weights, that is, $\sum_g N_g^a Y_g(a) / \sum_g N_g^a$ for both groups. In the figures, dots are the point estimates and bold lines are the third-order polynomial fits. Panel B shows the weighted averaged growth by age, that is, $\{\bar{\gamma}_g^a + \overline{\Delta\beta}_g\}$ where weights are again based on group counts. We also report the integrated incomes at age 25 and 55 from Panel A as we do in growth profile figures for subpopulations shown later.

As the figure shows, the growth profiles differ substantially across the groups. The self-employed incomes are significantly higher by age 55 than their paid-employed counterparts. Another striking difference is the life-cycle growth patterns shown in Panel B. Growth rates in paid-employment decline across the life cycle, while growth rates in self-employment do not. The attached self-employed have persistently high average growth in incomes—in the range of \$7 thousand to \$9 thousand annually—for ages between 25 and 40, whereas the paid have less than \$4 thousand in all years. Even after age 40, the self-employed average growth rate remains significantly higher than that for paid-employees and, at age 55, the self-employed have an average income of \$210 thousand—more than twice that of the paid-employed group and too large to be attributable to differences in taxes or benefits.

Growth Profiles for Attached Subgroups. Next, we repeat the exercise for subgroups of the attached self- and paid-employed, specifically by gender, marital status, education, skills, and industry. In Figure 3.3, we report growth profiles for men in Panel A and for women in Panel B. Recall that men account for most of the attached self-employment sample, roughly 82 percent, but only 53 percent of the attached paid-employment sample. Thus, it should not be a surprise that the self-employed growth pattern for men is nearly the same in magnitude and shape as the full sample. Perhaps more surprising is the growth pattern for self-employed women, which is also higher than for their paid-employed counterparts. For the paid-employed, neither the

men nor women show any increase in growth at age 25. When integrated, the income profiles reveal large level differences between paid- and self-employment for both men and women by age 55. However, we find a gender gap when we compare peak incomes, regardless of employment status. Average income for age 55 self-employed women is \$137 thousand—or 61 percent of that for men. Average income for age 55 paid-employed women is \$72 thousand—or 71 percent of that for men.

Another common attribute for the attached self-employed is being married in most years of the sample. Figure 3.4 reports growth profiles of the attached sample by marital status and incomes at age 25 and 55. Here again, we find similar results when comparing the mostly married to the full sample since they account for roughly 79 percent of the attached self-employed population and 68 percent of the attached paid-employed. The mostly unmarried self- and paid-employed have similar qualitative patterns in growth, but the incomes of the self-employed are \$58 thousand higher than that of the paid-employed by age 55.

Education is another important characteristic when considering returns to employment. After applying our classifier to categorize individuals as “likely” or “not likely” to have attained a college degree, we recompute slopes and construct average growth profiles. The results shown in Figure 3.5 reveal large differences in outcomes for those classified as college-educated when comparing across employment status and, in the case of the self-employed, when comparing across education levels. By age

55, we predict that the average income for college-educated self-employed individuals is \$321 thousand and only \$111 thousand for their paid-employed peers. The difference is even larger if we compare these self-employed individuals with those classified as not-college-educated. Average income at age 55 for the latter group is equal to \$42 thousand. Interestingly, as the figure shows, there is hardly any difference in outcomes if we compare self- to paid-employed if they are classified as not college-educated.

With the O*Net data and our occupation strings, we can study returns to self- or paid-employment for people with different skillsets. In Figure 3.6, we show growth profiles by cognitive skill. A comparison between this figure and that for education (Figure 3.5) reveals some difference in the categorizations. For example, differences for those classified as college-educated and not-college-educated are much starker than differences between cognitive and not-cognitive—regardless of occupation choice. In the case of the self-employed, those classified as cognitively skilled have an average income at age 55 of \$222 thousand while the not cognitively skilled have an average income of \$192 thousand.

In Figure 3.7, we show growth profiles by interpersonal skill. While not as stark as our comparisons by education, there are still large differences between the interpersonally skilled self- and paid-employed and between the skilled and unskilled regardless of employment status. For example, as with the college versus non-college

comparisons, we again find a much more pronounced hump-shaped growth profile for the interpersonally skilled self-employed when compared to the interpersonally unskilled self-employed. We again find that the non-interpersonally-skilled have much lower growth profiles and average incomes by age 55 when compared to their skilled peers, although not as low in magnitude as for the non-college-educated shown in Panel B of Figure 3.5. In Figure 3.8, we show growth profiles by manual skill. We find that patterns for individuals working in occupations with manual skills are similar to those described before for non-college group.

Another relevant cut of the data is by industry since the self-employed tend to be clustered in particular occupations and sectors. In Figure 3.9, we plot results for individuals with attached employment status that are in professional services (NAICS 54) and health care (NAICS 62). Together these sectors account for 27 percent of the attached self-employed population. Here, we see hump-shaped growth profiles for entrepreneurs and large differences between the income growth rates of the self- and paid-employed. By age 55, the attached self-employed average incomes in professional services and health care are \$304 thousand and \$283 thousand, respectively, whereas the average across all industries is \$210 thousand.

The results shown thus far are for relatively broad categories and include people with a wide range of characteristics. Because there are thousands of possible cuts of the data—even if we condition on being in the attached groups—plotting all of them

is not possible. However, we are able to highlight the most important groups by ranking them according to their importance in generating differences in the average income growth rate for the attached self-employed and the attached paid-employed. As shown in Figure 3.2, there is a sizable gap in income growth between ages 30 and 39—with an average at roughly \$5 thousand 2012 dollars. In Table 3.2, we summarize the groups that make up at least 50 percent of this difference. The first column summarizes the cumulative share. Reading across the row, we report distinguishing characteristics of the groups. We do not list the characteristics that they all share. It turns out that all top contributors are male, mostly married, college-educated, interpersonally skilled, and have kids. In terms of the distinguishing characteristics, we find that the group contributing the most to the growth differential works in the health care sector (NAICS 62), is born in the 1960s, is cognitively skilled, but is not manually skilled.

This group of primarily medical doctors contributes 10 percent to the growth gap, with the self-employed among them experiencing average annual growth in income around \$17 thousand during their 30s—roughly 1.6 times the growth of their paid-employed colleagues. As 3.2 makes clear, a small number of sectors matter for our results: health care, professional services, construction, finance, and retail trade.

Growth Profiles for Non-Attached Groups. Thus far, we have compared the attached self-employed to their paid-employed peers. Since a significant fraction of

self-employed income is earned by our mostly-switchers and any-non-employed sub-groups, we also include a comparison with these less-attached peers. In Figure 3.10, we compare the growth profiles for the latter groups to that of the attached self-employed. Panel A shows the growth profile comparison between the attached self-employed and the mostly switching. This figure highlights the delay in growth for those more attached to self-employment. Also interesting is the fact that the growth of the switchers is higher at age 25, but is declining in most years over the life cycle, reminiscent of the paid-employed growth patterns shown earlier.

Panel B shows the growth profile comparison between the attached self-employed and the any-non-employed.¹⁹ Individuals categorized as any-non-employed have little self-employment income on average but when aggregated, the sum is about 20 percent of all self-employment income. However, the growth profile for the any-non-employed have the same pattern as that of the paid-employed, just lower at all ages. This may not be too surprising given these individuals are less educated, lower-skilled, and likely to be second earners. But more surprising is the fact that many of these individuals are out of employment only one year and still have very low incomes and growth rates.

¹⁹When computing averages for the any-non-employed, we only include observations with non-zero incomes in both the previous and current year.

Residual Growth

We turn next to an investigation of the variability and persistence of income growth, which is relevant for discussions of entrepreneurial risk-taking and more generally for the larger debate on earnings inequality. We show that critical to this investigation is the estimated residual growth—that is, $\{\Delta\epsilon_{i,t}\}$, found after netting out the time and age effects from the total income changes—as this term accounts for most of the variation in overall growth. After reporting on key statistics of changes in overall and residual incomes, we work through a simple analytical example to demonstrate how these moments of the data can be used to estimate the potential gains of insuring income risk.

We start by computing year-to-year transitions in income *levels*, y_{it} , as a point of reference for both the self- and paid-employed. These matrices are reported in Table 3.3, with the results for the self-employed in Panel A and those for the paid-employed in Panel B. Each element of the matrix is the share of individuals that start the year in a particular income bin (listed in the first column) and transit to one of the income bins listed at the top of the matrix in the next year. Along the bottom of each matrix are the distributions of individuals in each bin. We have purposely chosen the positive income bins to have the same log-spacing between bins with 1 percent of the self-attached sample in the top bin. Earners in this bin have total incomes—from

self- and paid-employment—above \$1.6 million.²⁰ We also include a separate bin for negative incomes, which is possible for both the self- and paid-employed given we are reporting their total incomes that could include self-employment losses. Comparing rows across the two transitions, we find more dispersion for the self-employed in middle incomes when compared to the paid-employed, but less so at the top and bottom.

To provide a sense of the potentially risky nature of entrepreneurship across the life-cycle, we compute percentiles of growth rates by age, that is, $\Delta y_{i,a}/|y_{i,a-1}|$ between ages $a - 1$ and a for primary working ages 26 to 55 (and thus avoid later years when individuals are likely to work less or retire). Because incomes are both positive and negative, we divided the income changes by the absolute value of income in age $a - 1$. The percentiles of income growth are plotted in Figure 3.11, with results for the attached self-employed in Panel A and results for paid-employed in Panel B. For both groups, the income changes are most dispersed at younger ages. As expected, the self-employed incomes show more dispersion in growth rates at all ages. However, the 90–10 variation is relatively constant across middle ages for both groups, suggesting that the volatility in self-employed incomes is not rising over the life-cycle relative to

²⁰We also computed income transitions for the attached self-employed using deciles. DeBacker et al. [2022] report transition matrices with business income deciles but, given differences in sampling choices, their results are not directly comparable to ours. These authors only drop filers who *never* report business income outside of $[-5000, 5000]$ in 2012 dollars. Forty percent of the observations in their sample are included in bins below \$5,000: many more than we find for our attached self-employed sample. Furthermore, the lower cutoff for the top decile in their sample starts at \$82,000, whereas our cutoff when reporting deciles is \$333 thousand.

the volatility in paid-employed incomes despite the fact that differences in average incomes are rising over the life-cycle. This feature is relevant for theories that would rely on risk-compensation as an explanation for the differences in mean growth rates between paid- and self-employed individuals.

The data underlying Figure 3.11 can also be used to compute the variability of income changes within and across subgroups of our sample. Knowing this allows us to determine if there are significant differences in income risk faced by individuals in different subgroups, say for example, by those in different industries. There are two ways to compute the growth in observed incomes by age: we can pool the populations of attached self- and paid-employed and plot the dispersion as in Figure 3.11 or we can compute the statistics for subgroups of these populations and construct weighted averages using population counts for weights. Although not shown, we find that the dispersion in income changes is nearly the same regardless of whether we pool the individuals into attached self- or paid-employed or we instead compute a weighted average of the underlying subgroups. For both methods of aggregating, we find results similar to what is shown in Figure 3.11. This means that the within-group variation of income changes accounts for nearly all of the variation.

With time and age effects netted off, our estimation procedure yields residual growth for all individuals i in tax year t , namely, $\{\Delta\epsilon_{i,t}\}$, which is of independent interest when modeling shocks to income. As we show later, two features of these data

are particularly useful: the dispersion and autocorrelation. Figure 3.12 plots the percentiles of residual growth over the life-cycle after first normalizing the age $a - 1$ to age a change by the absolute value of lagged income ($|y_{i,a-1}|$). As we see from the figure, the results in Figure 3.12 reinforce the earlier evidence in Figure 3.11 and again show that the volatility in income changes is decreasing with age for both the self- and paid-employed. In Table 3.4, we compute the transition matrices for these data. As before, we report the share of individuals that start in a particular bin listed in the first column and transit to one of the income change bins listed at the top of the matrix in the next year. Distributions of these income changes are shown at the bottom of each matrix. These matrices can be used to infer how autocorrelated the income changes are and, in turn, how persistent or temporary they are. As we see from Panel A, the probabilities for the self-employed are high in the upper right and lower left of the matrix, indicative of a negatively autocorrelated process. The more negative the autocorrelation, the more temporary the income change. We see the same pattern for the paid-employed income changes reported in Panel B, which suggests that the main difference for the two groups is in the dispersion, not the persistence, of income changes.

The moments of the data discussed thus far can serve as useful inputs to welfare calculations of the gains to fully insuring against idiosyncratic risk. To demonstrate

how this is done, we work through a concrete example, making specific assumptions about the processes governing income growth—or, in our case, residual income growth—and the risk preferences of our individuals in order to derive analytical relations between data moments and welfare.²¹ Suppose that (i) growth rates in income are well summarized as the sum of a nonstationary random walk process r_t and a stationary autoregressive process z_t and (ii) preferences are of the [Epstein and Zin \[1989\]](#) class. To make analytical progress, we need a few more assumptions: the autoregressive process is not serially correlated, the shock processes are Gaussian, the intertemporal elasticity of substitution parameter in preferences is equal to 1, and consumption moves one-for-one with permanent shocks to income and not at all to transitory shocks as dictated by the permanent income hypothesis. Then, with simple algebra, it is straightforward to show that the fraction of wealth (λ) an individual would forgo to fully insure their risky income is given by:

$$\lambda = -\frac{1}{2}\alpha\beta\sigma_r^2, \tag{3.7}$$

where σ_r^2 is the variance of the Gaussian shocks of the random walk process and α and β are parameter inputs to the utility function V over paths of consumption $\{c_j\}$

²¹With the aid of computer simulations, it is straightforward to relax these assumptions and consider more general specifications.

proposed by [Epstein and Zin \[1989\]](#), namely,

$$V_t(\{c_j\}_{j=t}^{\infty}) = \left[(1 - \beta)c_t^{\rho} + \beta(E_t V_{t+1}^{\alpha})^{\rho/\alpha} \right]^{\frac{1}{\rho}}. \quad (3.8)$$

Equation (3.7) is derived by taking the limit as ρ approaches 0, consistent with an intertemporal elasticity of substitution equal to 1.

To make progress quantitatively, we can use estimates of 90–10 differences and autocorrelations for income changes to infer the variance σ_r^2 in (3.7). With these moments and the assumptions given above, we can derive an analytical relation between the IRS statistics and predictions for variances of the permanent and temporary shocks underlying the income changes. Let Q be the 90–10 difference in the income changes and let A be the autocorrelation. Then, we can show:

$$Q = 2.56\sqrt{\sigma_r^2 + 2\sigma_z^2} \quad (3.9)$$

$$A = -\frac{\sigma_z^2}{\sigma_r^2 + 2\sigma_z^2}, \quad (3.10)$$

where the 2.56 is equal to the 90–10 difference for a standard normal and $\sigma_r^2 + 2\sigma_z^2$ is the total variance of the income change if the temporary shock z_t is an independent and identically distributed random variable.²² Note that the variance of the temporary shock is multiplied by two since we are analyzing the change in income. From

²²We derive these results along with equation (3.7) in a separate appendix.

equations (3.9)-(3.10), we derive an estimate of the variance for this example:

$$\sigma_r^2 = \left(\frac{Q}{2.56} \right)^2 (1 + 2A). \quad (3.11)$$

In the case of the residual growth process, estimates for the 90–10 differences can be read right off of Figure 3.12. If we average over the life-cycle, we find Q around 1.61. Using Table 3.4 for the self-employed, we find A to be roughly -0.16. Using the formula above, we then have an estimate for the permanent shock standard deviation σ_r of 0.52 and an estimate of the temporary shock standard deviation σ_z of 0.25. Suppose we set the discount factor β in the Epstein-Zin preferences equal to 0.96 and the risk aversion parameter α equal to -1. In this case, we would predict that the fraction of wealth (λ) an individual would forgo to fully insure their risky income is 12.8 percent. If we repeat the exercise for the paid-employed and assume the same preference parameters, we find that λ is around 1.4 percent. In this case, differences in λ are driven primarily by differences in dispersion (which has a ratio around 3) because the autocorrelations are not that different between the two groups (-0.16 versus -0.17).²³

While the exercise of translating the data moments into measures like λ provides

²³Abowd and Card [1989] report an average autocorrelation of -0.29 for changes in logged earnings of paid-employed male household heads in the Panel Survey of Income Dynamics (excluding the Survey of Economic Opportunity subsample), with the range of estimates for 1969–1979 given by $[-0.54, -0.10]$. We find an average autocorrelation of $\Delta \log y_{it}$ for the paid-employed equal to -0.25 , regardless of whether we restrict attention to males or married males or include all paid-employed.

economically interpretable summaries of the risk entrepreneurs face, it is important to keep in mind that these calculations assume a one-for-one pass-through of income shocks to consumption, which allows analytical tractability but is too extreme. We know from Table 3.1 that 79 percent of these individuals are married most years of our sample. As we report at the end of Table 3.1, the average wage of their spouses is \$26 thousand before accounting for any employee benefits. We also know that the average household asset income for this group—including payments of dividends, interest, and capital gains—is \$58 thousand, implying sufficient saving stocks in downturns.

More generally, if the total adjusted gross income (AGI) is a better proxy of what household members have available for consumption, then a more quantitatively relevant pass-through rate is the change in AGI following a change in self-employment income. To compute this, we estimate equation (3.2) with y_{it} set equal to AGI and construct new residual growth rates. We then regress these rates on those computed above (with y_{it} equal to income of the attached self-employed individuals).²⁴ When we do this, we find a coefficient from the regression equal to 0.16—well below 1 assumed in the calculations above.²⁵ In other words, the analytical calculations above, while economically interpretable, should be viewed as an upper bound on the welfare

²⁴Before running the regression, we drop the top and bottom 0.01 percent outliers for both residual growth rates.

²⁵Entrepreneurs also have outside opportunities that provide additional insurance such as switching to paid-employment, declaring bankruptcy, or listing their firms. See, for example, Manso [2016] and Catherine [2022].

gains for fully insuring consumption as individuals in our sample already partially do through other means.

3.4.2 Entrepreneurial Choice

We turn next to analyzing individuals who switch across employment status. Understanding switching behavior is key for theories of occupational choice. In this section, we measure the extent of switching and analyze differences between those who switch and those who do not. We find that entry rates into self-employment are relatively flat across the life cycle and across time. Exit rates out of self-employment decline with age but vary little over time, even during the 2008–2009 recessionary period. Importantly for theory, we find that the switching behavior reveals positive selection on past incomes, negative selection on asset and spousal wage income, and roles for both non-pecuniary and pecuniary motives driving entry in and exit out of self-employment.

We focus on switching rates between employment activities, namely, paid-, self-, and non-employment. For a group of individuals, a switching rate from activity A to B is defined as the fraction of individuals whose status was A at age a (or date t) and B at age $a + 1$ (or date $t + 1$). The entry rate into activity A is the fraction of individuals who transit from not- A at age a (or date t) to A at age $a + 1$ (or date $t + 1$), and the exit rate is defined analogously.

In Panel A of Figure 3.13, we plot the entry rates into self-employment from either paid- or non-employment or both by age. The figure shows that the overall entry rate is in the range of 1 percent to 2 percent and is modestly increasing in age. Although not shown here, there is also a distinct gender gap in the overall entry rate, with women's rates significantly lower. As is clear from the figure, most of the rise is due to entry from non-employment. If we condition on gender, we find that most entry from non-employment is actually men and not simply a return of women to the ranks of the employed after having children.

Exit rates out of self-employment are shown in Panel B. The overall rate is high and strongly declining, starting around 38 percent and dropping to about 17 percent by the end. If we compute these statistics for women and men, we again see a gender gap: relative to men, women have exit rates that are roughly 6 percent higher at all ages. The declining hazard rate—whether for men or women—is suggestive that experimentation and learning about the potential gains to entrepreneurship occurs early in careers. Most of those switching at early ages go into paid-employment. Not surprisingly, by the end of the life cycle, more switch to non-employment because of early retirements.

In Figure 3.14, we plot the entry and exit rates by tax year. In this case, we purge an age effect that arises due to the aging population over our sample period.²⁶ We

²⁶Specifically, after constructing switching rates for each date t , we subtract a weighted difference of the age- a switching rates, with weights equal to the age- a share of the population in t less the age- a share of the population in year 2001.

find that, between 2001 and 2015, the entry and exit rates are remarkably flat with no clear time trend. The lack of cyclical variation around 2008–2009 suggests that self-employment is not used by many as a hedge against unemployment risk. (See, for example, [Alba-Ramirez \[1994\]](#), [Evans and Leighton \[1989\]](#), [Rissman \[2003\]](#), and [Rissman \[2007\]](#).)

To better understand the motives and impediments to switching, we compare past incomes for individuals who switch at some age to that of comparable individuals who switch at some later age.²⁷ Our first exercise compares past wage income for one-time switchers into self-employment to past wage income of those who switch later, but share the birth year, gender, industry, marital status, and lagged employment status, whether it be paid- or non-employment.²⁸ We use three years of data when computing past incomes and when comparing lagged employment status. More specifically, let $x_{i,t}$ be the variable used to predict the switch—say, past wage income in this case—for individual i at time t and let $x_{m(i),t}$ be the same variable for all matched peers $m(i)$ of individual i . Then we compute the difference Δ_{it} in the averages of variable x before the switch as

$$\Delta_{it} = \frac{1}{3} \left[\sum_{j=1}^3 x_{i,t-j} - \frac{1}{N_{m(i)}} \sum_{m(i)} \sum_{j=1}^3 x_{m(i),t-j} \right]. \quad (3.12)$$

²⁷We also compared results for switchers to that of comparable non-switchers and find quantitatively similar results for all panels of Figure 3.15.

²⁸In these comparisons, we use yearly indicators of married or not married rather than the time-invariant notion of “mostly” married.

In Panel A of Figure 3.15, we plot the interquartiles of this difference in past income by age of switch. A positive value indicates switchers have higher past income than the future switchers. We see that early switchers have similar past incomes to non-switchers and over time the gap becomes larger and more favorable for the switchers. By age 55, the median difference is roughly \$2 thousand. These findings hold up even if we focus exclusively on those in paid-employment prior to the switch. From this exercise, we conclude that most switchers are positively selected on past productivity. Next, we compare asset incomes of switchers and future switchers.²⁹ Since labor and asset incomes are generally correlated, we isolate the role of assets for entrepreneurial choice by comparing switchers to later-switching peers who not only share birth year, gender, industry, marital status, and lagged employment status, but also the percentile of past wage income. In Panel B of Figure 3.15, we plot the distribution over age of past asset income of the switchers less an average of past asset income of the future switchers they are paired with. Call this difference the excess asset income. For most switchers, we find the differences to be negative and small. As in the case of the past wage income comparisons, the differences in incomes are small for younger ages and grow larger (in absolute value) for older ages. By age 55, the median difference in asset incomes is $-\$2$, with the current switchers earning less than the future switchers. From this exercise, we conclude that most switchers are negatively selected on liquidity.

²⁹For these results, we exclude Schedule D capital gains from asset income.

Panels C and D of Figure 3.15 repeat the exercise but in these panels, we use past spousal wages and adjusted gross incomes instead of asset incomes. Like asset incomes, other incomes earned by the household can be used for financing business startups or for smoothing consumption in the early years of operations. In the case of spousal wages, we find the differences to be small or negative for most of the distribution. Not surprisingly, in the case of adjusted gross income—which is a much broader measure of income—we find larger differences across the distribution when comparing switchers and future switchers. However, the differences here are roughly split between positive and negative and, thus, there is no definitive conclusion about selection.

In order to investigate pecuniary versus non-pecuniary motives for switching, we compute the change in income following a switch in employment status from self- to paid-employment or vice versa. Declines in income would be consistent with a potential role for non-pecuniary motives. To test this, we compare average incomes of individuals that switch their employment status—averaging over the incomes in the three years subsequent to the switch—and compare this to the 3-year averaged income of a control group that did not switch. As above, we assume the control group has the same birth year, gender, industry, marital status, and 3-year lagged employment status as the switcher, but remains in the same employment status for at least three more years. We also condition on the skillset—which is a good proxy

for occupation—and the fact that individuals in the control group do switch their status at some point outside of the six-year window. In Panel A of Figure 3.16, we plot the income differential following a switch—along with the interquartile ranges—weighted by subgroup counts, for ages 28 to 60. We find the median income change post-switch is negative but small. Overall, we find a large fraction of the population of switchers earning more post-switch and a large fraction earning less than non-switching peers, suggesting that there are both non-pecuniary and pecuniary motives driving entrepreneurship.

Results for those switching from self- to paid-employment are shown in Panel B of Figure 3.16. Here again, we find that the results are split: there are almost as many with higher post-switch income as there are with lower. The exceptions are the youngest entrepreneurs that see more pay gains than losses when switching to paid-employment. We turn to examining this group in more depth next.

3.4.3 Young Entrepreneurs

Our results thus far suggest that entrepreneurs under the age of 40 have higher exit rates when compared to older peers and, at least for those that stay, have steeper growth profiles and more volatile incomes. This subgroup of our sample is particularly interesting because young business owners are less likely than older peers to have previous self-employment experiences or opportunities to accumulate much

in the way of assets before starting a business. With this in mind, we revisit some themes already discussed above but with a narrower focus on the youngest cohorts born between 1970 and 1975.

To investigate the high exit rates of young entrepreneurs, we ask if there are differences between young individuals who experiment with self-employment while young and continue on in business with those that experiment but then exit. We track individuals in the youngest cohorts with at least five years of self-employment experience prior to age 35. In Figure 3.17, we report the growth profiles for those that continue in self-employment after age 35 to those that switch to paid-employment. The figure reveals a familiar pattern: the growth profile of those continuing in self-employment is higher and more hump-shaped than the profile of those who switch into paid-employment. Those that stay have profiles more similar to attached self-employed and those that exit have profiles more similar to attach paid-employed. One explanation is that the switchers were never committed to the entrepreneurial path in the first place and did not make the necessary firm-specific investments. Another explanation is that the switchers learned early that they have low entrepreneurial skill and exited. Later, we test these hypotheses.

In terms of volatility, a potential issue for young entrepreneurs—especially those starting new businesses—is financing start-up costs and early firm-specific investments. Here, we characterize the path of initial losses at both the business level and

individual level for a group of business founders that are in one of the 1970–1975 cohorts. We define founders to be individuals filing a Schedule K-1 as an owner in the first year that a S-corporation or partnership starts. We further restrict attention to businesses that have at least eight years of consecutive tax filings with business receipts or deductions. For this sample, we find that the business net income is negative for 45 percent in the first year of business, 35 percent in the second, and 32 percent in the third. Flipping this around, we ask when these businesses have a first positive net income. In this case, we find 53 percent of businesses in the first year, 19 percent in the second, and 8 percent by the third.

Interestingly, if we ask these questions of the new owners in terms of their own self-plus paid-employment income, we find a relatively low pass-through of losses. In the first three years of business, we find that the owner’s total income (from self- plus paid-employment) is negative for 10 percent of owners in the first year, 9 percent in the second, and 8 percent in the third. If we ask when these owners have their first positive total income, we find 90 percent in the first year, 5 percent in the second, and 2 percent in the third. The higher share of owners with positive income at an earlier stage is further evidence that self-employed individuals have other means of smoothing personal consumption expenditures over time—say, because they have multiple businesses or because they supplement the early year business incomes with income from paid-employment.

3.4.4 Comparison to existing literature

To better motivate the theory that we develop next, we first relate our empirical findings to those in the existing literature. There is a large literature that uses survey data for the United States to investigate entrepreneurial income profiles and occupational choice. Prominent examples are [Lazear and Moore \[1984\]](#) with the Current Population Survey, [Evans and Leighton \[1989\]](#) with National Longitudinal Survey of Youth (NLSY), [Hamilton \[2000\]](#) with Survey of Income and Program Participation, [Hurst and Lusardi \[2004\]](#) with the Panel Study of Income Dynamics, and [Moskowitz and Vissing-Jorgensen \[2002\]](#) and [Kartashova \[2014\]](#) with the Survey of Consumer Finances.³⁰ This literature has been extremely influential in promulgating our understanding of entrepreneurship and motivating theories that can be used for policy analysis. In this section, we relate our findings to these studies—delineating points of agreement and points of disagreement.

Since survey data have issues related to top-coding and small samples, most research on entrepreneurship has focused on the median incomes of the self-employed. To relate our findings to those based on surveys, we start by comparing simple cross-sectional moments—medians and means—for self-employment income and paid-employment-income based on data from the CPS and IRS. The IRS sample used for this comparison is different than that summarized in Table [3.1](#) in order to ensure consistency with

³⁰For a comprehensive set of references, see [Parker \[2018\]](#).

the CPS. More specifically, we use data for all available cohorts and only two criteria for the self-employment assignment: the absolute value of income must exceed \$5000 (in 2012 U.S. dollars) and must be greater than the income from paid-employment.³¹ In Panel A of Figure 3.18, we plot median self-employment incomes by age for the IRS and CPS.³² While the IRS profile is lower and steeper at early ages than that based on CPS data, the series show median incomes that are close between the ages 45 and 60. If we compare either series to the paid-employed counterparts in Panel B of Figure 3.18—where we find almost no difference between CPS and IRS before age 55—we recover a familiar result: median self-employment income is below median paid-employment income. These results are consistent with an abundance of survey evidence that finds a self-employment “discount.”³³ This finding has solidified a view that self-employed individuals must be earning large non-pecuniary benefits from being their own boss and having flexible jobs. (See, for example, [Hurst and Pugsley \[2011\]](#) and [Catherine \[2022\]](#).) Similar conclusions are drawn by [Moskowitz and Vissing-Jorgensen \[2002\]](#) who emphasize low returns relative to the risk in self-employment.³⁴

³¹As before, we exclude the top and bottom 0.01 percent outliers. Including these individuals adds more noise to the series of cross-sectional means but barely changes the polynomial fit.

³²Quantitative results shown in Figure 3.18 are robust to using the main sample of analysis underlying Table 3.1.

³³Although [Levine and Rubinstein \[2016\]](#) claim that median incomes are higher for the incorporated self-employed, their estimate of the difference is only \$5 thousand above the paid-employed for the NLSY. This finding is consistent with [Hamilton \[2000\]](#), who documents smaller differences between incomes of the self- and paid-employed at higher quantiles of the distribution where the incorporated owners would naturally be.

³⁴[Hall and Woodward \[2010\]](#) analyze data for the universe of U.S.-based high-tech startups and

In the lower panels of Figure 3.18, we plot cross-sectional means for each age. Panel C has results for self-employment incomes reported in the CPS and the full IRS sample and Panel D has analogous results for paid-employment incomes. There are two comparisons to be made here. First, we need to compare survey results versus the IRS population. In the case of self-employment income, the differences are large: the IRS average income is close to \$42 thousand higher than the CPS average at the peak.³⁵ In the case of paid-employment, the differences are small.³⁶ The second comparison is across employment status. Using the IRS data, we would conclude that the self-employed earn significantly more than the paid-employed. Using the CPS data, we would conclude that the paid-employed—in particular, those in prime working ages—earn only modestly more than the self-employed.³⁷

The CPS-IRS comparisons across the means and medians suggest that the discrepancies are driven by the properties of the right tail. To investigate this further, we compare the 75th, 90th, 95th percentiles and shares of incomes above those

find expected returns for the owners are higher than from salaried work, but also conclude that the returns are roughly equal after taking into account the risk owners face.

³⁵This difference does not correct potential underreporting in the tax data. For instance, for a matched CPS-IRS sample of self-employed individuals, ? find that CPS respondents report 51 percent *less* income to the IRS.

³⁶Bollinger et al. [2019] compare average CPS and SSA W-2 earnings for 440 thousand individuals that appear in both samples. Over the period 2005–2010, they find the difference in average income is \$813 (reported in 2010 dollars).

³⁷The econometric approach laid out in Section 3.3.2 exploits the long panel that administrative data provides to separately analyze groups that differ with respect to their attachment to self-employment. This is the main advantage over working with simple averages.

percentiles in comparable CPS-IRS groups. For self-employed individuals, the differences are stark. The right-tails in the CPS are significantly thinner than their IRS counterparts.

The fact that the self-employment income distribution is right-skewed means that the typical dollar in self-employment does not come from the typical self-employed individual. To explore this further, we introduce a new group based on ranking individuals on their average income, once we have conditioned on their NAICS code, cohort, and gender. To ensure that we compare slopes of income profiles by employment status for individuals with similar average incomes, we deliberately ignore their employment status (paid- versus self-employment) before assigning them a rank. After ranking them, we bin individuals into five quintiles. Table 3.5 shows the shares of income after individuals have been ranked for total, paid-, and self-employed income. In the case of self-employed income, we see that 89 percent of the self-employed income comes from individuals in the top 25 percentiles and a majority of the latter are those we classified as attached self-employed.

In Figure 3.19, we plot the growth profile for those in the top 25 ranks in the case of the attached self- and paid-employed (Panel A) and those in the bottom 25 ranks (Panel B). The figures show that for the top 25 ranks the differences in growth rate by age are starkly different across self- and paid-employed while for the bottom 25 ranks they are not. From this, our take away from this is that most of self-employment

income is characterized by the patterns we highlighted in Section 3.4.1, that is, with steeper, more persistent, income growth for the self-employed as compared to the paid-employed. In other words, we find patterns for the top ranks that are quite different from those emphasized by the current literature. For the bottom ranks, on the other hand, we see the self- and paid-employed patterns are similar, as we saw earlier in subgroups like the non-college-educated. However, these groups have only a small share of either paid- or self-employment income.

Next, we compare our findings on entrepreneurial entry and exit. As far as switching rates by age are concerned, our estimates are in line with those from surveys. (See, for example, [Evans and Leighton \[1989\]](#) and [Fairlie \[2005\]](#)). Our findings that entry and exit rates do not show a trend or fluctuate much around the 2008–2009 recession might seem contradictory to the findings from U.S. Census data, such as [Decker et al. \[2014\]](#), which shows a decline in the startup rate around 25 percent at the start of the Great Recession. However, their findings largely reflect differences in samples. Studies that find declining entry rates use measures such as the fraction of new firms in the Longitudinal Business Dynamics (LBD) data, while we focus on individuals who enter self-employment. To reconcile the differences, we analyze a smaller sample of self-employed that is more aligned with the LBD, namely, those with employees. This group includes about one-third of self-employed individuals and account for two-thirds of self-employment income. If we recompute the change in entry rates for

this subgroup of self-employed, we find declines that are in line with [Decker et al. \[2014\]](#) and [Bayard et al. \[2018\]](#).

Where we differ with the literature is in our conclusions concerning selection into entrepreneurship. A common finding from previous work is that individuals entering self-employment have lower past labor incomes when compared to peers that are similar but did not enter. As [Evans and Leighton \[1989\]](#) explain, such findings are consistent with sociological views that “misfits,” who are poorer wage earners and more likely to change jobs, are more likely to be self-employed.³⁸ This is contrary to our findings, which show that most individuals entering self-employment have *higher* past labor incomes relative to similar peers that did not enter. (See Figure [3.15](#).) Since we also find that there are strong pecuniary motives for many of our switchers and large ex-post returns to choosing self-employment, there may be top-coding issues with the survey data that are leading to differences in our conclusions.

Another common finding is that individuals that enter self-employment have greater holdings of financial assets. This finding has sparked a large literature emphasizing significant liquidity requirements as impediments to self-employment. See, for example, the work of [Evans and Jovanovic \[1989\]](#), [Quadrini \[1999\]](#), [Cagetti and DeNardi \[2006\]](#), and [Buera \[2009\]](#). There are notable exceptions, namely, [Hurst and Lusardi \[2004\]](#) and [Fairlie \[2005\]](#), who find a limited role for liquid assets as determinants of self-employment. As we showed earlier, when we compare self-employment entrants

³⁸See also [Alba-Ramirez \[1994\]](#), [Rissman \[2003\]](#), and [Rissman \[2007\]](#).

to comparable non-entrants, we find the latter has higher average asset income. Thus, we view our findings as strengthening the conclusions of [Hurst and Lusardi \[2004\]](#) and [Fairlie \[2005\]](#).

3.5 Implications for Theory

From our empirical analysis, we find several salient empirical patterns that can be used to inform theories of entrepreneurship. The attached self-employed have persistently high income growth profiles as compared to paid-employed peers. Exit rates are high when individuals are young and decline steeply over the life cycle. Volatility of income changes is higher in self- than paid-employment but does not rise with age. In this section, we analyze a theoretical model that incorporates features motivated by these empirical findings and use it to determine how well the predictions for incomes and growth align with empirical counterparts. Here, we focus on the decision-making of young entrepreneurs—those that are attached to their employment status and those that ultimately switch—but the model is sufficiently general to be used in future analyses to study other aspects of the data described earlier.

In the spirit of [Bhandari and McGrattan \[2021\]](#), we incorporate firm-specific investments in self-created intangible assets—customer bases, client lists, inventions, designs, processes—that are needed before production can begin at an optimal scale.

In the spirit of Jovanovic [1982], we assume returns on these investments are uncertain because our founders have no previous experience and must learn about their productive capabilities for running a business. As they gain experience, they choose to continue with the business or to discontinue, selling their intangible assets and switching to paid-employment following the exit. When or if they exit depends on the productivity shocks these owners face in self-employment and in paid-employment if they do switch.

3.5.1 Model

Each period, entrepreneurs decide if they will continue running their businesses, sell them, or discontinue without sale and then work for someone else. They condition these decisions on the state s , which depends on financial asset holdings a , business intangible assets κ , ability in paid-employment ϵ , latent ability in self-employment z , and years of experience j . Because z is latent, the predicted mean μ of ability, which depends on past observations of productivity, is also included in the state vector s . Owners that decide to keep the business choose consumption and inputs for intangible investment and production. Intangible investment requires time, h_κ , and expenses, e , which are inputs in the technology $f_\kappa(h_\kappa, e)$. Production requires the stock of intangible assets, κ , hours of the business owner, h_y , and external factors that can be rented, namely, tangible capital, k , and labor n , which are inputs in the technology

$f_y(\kappa, h_y, k, n)$. The prices for the external capital and labor are r and w , respectively, and taken as given by the businesses. Denoting the value of keeping the business by $V_k(s)$, we formulate the problem as a dynamic program:

$$\begin{aligned}
V_k(s) &= \max_{c, h_y, h_\kappa, k, n, e} \{U(c, \ell) + \beta EV(s')\} & (3.13) \\
\text{subject to } a' &= (1+r)a + pe^z f_y(\kappa, h_y, k, n) - (r + \delta_k)k - wn - e - c \\
\kappa' &= (1 - \delta_\kappa)\kappa + f_\kappa(h_\kappa, e) \\
\ell &= 1 - h_y - h_\kappa \\
a' &\geq 0
\end{aligned}$$

and additionally processes for updating ϵ , z , and μ described below. Goods and services sold by the business have a unit price of p and the capital stocks are assumed to depreciate at rate δ_k for tangible capital and δ_κ for intangible capital. In this problem, the value next period is $V(s')$ and is the maximum value of the three alternatives: continuing $V_k(s')$, discontinuing with sale $V_s(s')$, and discontinuing without sale $V_w(s')$. The value is a discounted sum of period utilities $U(c, \ell)$ defined over consumption and leisure, with future values discounted by factor $\beta \in [0, 1)$.

The entrepreneurial productivity has a permanent unobserved component \bar{z}_0 plus an idiosyncratic and identically distributed shock η_j . Let z_j denote the value of

productivity for an entrepreneur with j years of experience. In this case,

$$z_j = \bar{z}_0 + \eta_j$$

and we assume that η_j is normally distributed with mean 0 and variance σ_η^2 that is constant for all j . Let μ_j and σ_j^2 denote the predicted mean and variance of the latent productivity conditioned on past signals, that is:

$$\begin{aligned}\mu_j &= E[\bar{z}_0 | z_0, z_1, \dots, z_{j-1}] \\ \sigma_j^2 &= \text{var}[\bar{z}_0 | z_0, z_1, \dots, z_{j-1}],\end{aligned}$$

and the distribution of z_j conditional on the past observations z_0, z_1, \dots, z_{j-1} is assumed to be normally distributed with mean μ_j and variance $\sigma_j^2 + \sigma_\eta^2$. In this case, the means and variances evolve with one more year of experience as follows:

$$\begin{aligned}\mu_{j+1} &= \mu_j + \frac{\sigma_j^2}{\sigma_j^2 + \sigma_\eta^2}(z_j - \mu_j) \\ \sigma_{j+1}^2 &= \frac{\sigma_j^2 \sigma_\eta^2}{\sigma_j^2 + \sigma_\eta^2}.\end{aligned}$$

The next period value in (3.13) also depends on the evolution of the paid-employment ability because the entrepreneur can decide to sell or discontinue with business. For

tractability, we assume that entrepreneurs that sell or choose to switch to paid-employment do not start a business after that. Relevant to this choice is the evolution of ϵ which is Markov with transition probability $\pi(\epsilon'|\epsilon)$. Then the value of sale is given by

$$\begin{aligned}
V_s(s) &= \max_{c, h_y, k, n} \{U(c, \ell) + \beta \sum_{\epsilon'|\epsilon} \pi(\epsilon'|\epsilon) V_w(s')\} & (3.14) \\
\text{subject to } a' &= (1+r)a + pe^z f_y(\kappa, h_y, k, n) - (r + \delta_k)k - wn + p_\kappa \kappa - c \\
\ell &= 1 - h_y
\end{aligned}$$

with $a' \geq 0$ and $\kappa' = 0$ and revenues from the sale given by $p_\kappa \kappa$. The value of working for others, V_w , is standard and given by

$$\begin{aligned}
V_w(s) &= \max_{c, h} \{U(c, \ell) + \beta \sum_{\epsilon'|\epsilon} \pi(\epsilon'|\epsilon) V_w(s')\} & (3.15) \\
\text{subject to } a' &= (1+r)a + w\epsilon h_y - c \\
\ell &= 1 - h_y
\end{aligned}$$

with $a' \geq 0$, where again for tractability, we have assumed that paid employees do not switch to self-employment mid-career.

3.5.2 Quantitative results

Next, we analyze numerical simulations of the entrepreneurial optimization problem and compare predicted growth profiles with empirical counterparts. For our baseline parameterization, we use estimates for preferences and technologies from [Bhandari and McGrattan \[2021\]](#) based on aggregate data from the Bureau of Economic Analysis’s (BEA) national accounts and micro data from the Census’s Survey of Business Owners (SBO) and the Pratt’s Stats database of brokered business sales.³⁹ We use the IRS micro data to ensure that the variability and persistence of incomes are consistent across model and data.

The functional forms for preferences and technologies used by [Bhandari and McGrattan \[2021\]](#) are given by:

$$\begin{aligned} U(c, \ell) &= (c^{1-\psi} \ell^\psi)^{1-\sigma} / (1-\sigma) \\ f_\kappa(h_\kappa, e) &= h_\kappa^\vartheta e^{1-\vartheta} \\ f_y(\kappa, h_y, k, n) &= \kappa^\phi k^\alpha (\omega h_y^\rho + (1-\omega)n^\rho)^{\frac{\nu}{\rho}}, \end{aligned}$$

with values for parameters listed in Table 3.6. There are three parameters related to preferences: ψ , σ , and β . Setting the weight on leisure, ψ , to 58 percent ensures that levels of aggregated business hours are consistent with U.S. totals. The value of 1.5

³⁹Estimating the full model requires linking all business filings for the self-employed and is beyond the scope of this paper.

for σ is standard in the literature. The value of 0.96 for β is consistent with U.S. real returns to capital of roughly 4 percent. In terms of technology parameters, the most relevant for the income and growth profiles is the share of intangible capital in the production of goods and services, ϕ . The share ϕ affects founders' incentives to invest time and resources in building their business. If the revenue share is small and only external factors are required, then growth of the productive self-employed will be high relative to paid-employed or entrepreneurs that switch out of self employment early. [Bhandari and McGrattan \[2021\]](#) jointly estimated this parameter, along with two other parameters governing intangible production (ϑ, ρ) using information about the intangible share of assets in business sales, the input shares from the BEA input-output tables, and the entry rate of new businesses. They reported estimates of $\phi = 0.15$, $\vartheta = 0.408$, and $\rho = 0.500$. The remaining production shares, namely, $\alpha = 0.3$, $\omega = 0.425$, and $\nu = 0.55$, are based on revenue shares in U.S. private business data. Finally, the depreciation rates used by [Bhandari and McGrattan \[2021\]](#) are based on studies of depreciable and amortizable assets conducted by the BEA and IRS and set equal to $\delta_k = 4.1$ percent and $\delta_\kappa = 5.8$ percent.

IRS micro data are used to estimate prices and productivities and reported in the lower panels of Table [3.6](#). Two prices are pre-set: the interest rate at $r = 4.1$ percent, which is consistent with the preference parameters, and the wage rate at $w = 1$, which

is a normalization. Given this value of the wage, we set the price of goods and services p equal to 1.5 to ensure that the relative income for the young entrepreneurs—stayers versus switchers—is consistent with IRS data at age 40. The price per unit of self-created capital, p_κ , if sold is set equal to 1.6 and chosen to generate ratios of the business value to seller’s wage bill between 2 and 3, consistent with U.S. private business sales. For entrepreneurial productivities, we normalize the predicted mean μ_0 to 0 (along with the average \bar{z}_0). Shock variances driving entrepreneurial income shocks are chosen to ensure that the 90–10 differences and autocorrelations in business income changes are the same in the model and the IRS data. This implies values for the predicted initial productivity variance, σ_0^2 , equal to 0.5 percent and the idiosyncratic variance, σ_η^2 , equal to 0.4 percent. Similarly, the parameters governing paid-employment income are set to ensure that the 90–10 differences and autocorrelations of employee income changes—relevant for those exiting to paid-employment—are the same in the model and the IRS data. To model $\pi(\epsilon'|\epsilon)$, we use the method of [Tauchen \[1986\]](#) to approximate a continuous autoregressive process as a Markov Chain. When simulating our data, we use 11 states, a persistence parameter of 0.7 and a standard deviation of 0.1.

Given parameter estimates, we now use our laboratory to simulate income and growth profiles for a large sample of entrepreneurs. In the simulations, we assume a full life cycle of 60 years but report results for ages 25 to 40 in order to compare our

predictions to the data on young entrepreneurs—both the stayers and switchers. We assume that these young startups have no assets and little in the way of transfers—just 0.01 so that initial consumption is not zero. As we vary these choices, we find that there is only a small impact on production decisions for entrepreneurs with high predicted productivity levels, who want to scale up as quickly as possible. To make the model and data results comparable, we use the counts by employment status from the 1975 cohort, which includes individuals between 25 and 40 during our sample period. For example, we know how many are self-employed at age 25, how many at age 26, and so on. Using these counts, we find a roughly constant entry rate into self-employment between ages 25 and 31, with rates on the order of 11 percent per year. Using this constant rate, we extrapolate back to age 22, which is before we see them in the sample. To compute theoretical predictions, we simulate data for 22-year olds, 23-year olds, and so on, and then use the actual counts of self-, paid-, and non-employed to weight the model-generated incomes (which is equal to the average wage for paid-employed and zero for non-employed). Then, we construct income and growth profiles for two groups from the model simulations: self-employed stayers and self-employed switchers. Both groups have at least five years of either self- or paid-employment experience prior to age 35. After 35, the self-employed stayers have continued on in self-employment and the self-employed switchers have discontinued or sold their business and switched to paid-employment. Weights from

the 1975 cohort are then used to add up the stayers and the switchers at different ages.

In Figure 3.20, we show the differenced income growth from the data (that is, differences in the two profiles in Figure 3.17 against the predicted profile. We should note that we have not included any economy-wide technological changes or alternative sources of growth that would be common to individuals with different employment status. In fact, in the model, the outside opportunity of paid employment is a flat income profile when averaged. Therefore, we compute differences in growth profiles for both the model and data so they are comparable. Both show a humped growth profile. In the model, we generate this because of two key features: learning and investment. If we abstract from either, then we are unable to generate this pattern. Consider first the role of learning. At the start of the simulations, the self-employed that ultimately stay or switch look the same. They all start with an initial prediction μ_0 of 0 and the same variance on the productivity signal. After that, they gain experience and those that ultimately exit self-employment have a mean prediction for their productivity that has fallen over time. Importantly, the fall in the mean prediction leads these entrepreneurs to reduce their investments in self-created intangibles over time. Less investment means less growth in subsequent years and thus, an eventual exit to better opportunities in paid-employment. If there is greater certainty about the entrepreneurs productive capability—that is, if the variance σ_η^2

is lowered—exits occur earlier. If it is sufficiently low, then exits occur immediately and we would not observe any entrepreneurs waiting five years before switching to paid-employment.

The second key feature is the investment in κ made by entrepreneurs. For our baseline parameters, we find that roughly 10 percent of available time is used initially to build κ . By age 35, investment is close to zero for the switchers but around 10 percent for the stayers. Noteworthy is the fact that entrepreneurs who ultimately stay in self-employment start increasing their investments immediately in order to quickly build up their intangible capital stocks. These investments ultimately pay off in higher incomes later. As they build the intangible stock, entrepreneurs start to substitute external hours from paid employees for their own time in production of goods and services. For the entrepreneurs that continue past the age of 35, we find a steady drop in own hours of production and a scaling up of the business as they hire external labor and capital. By age 40, the ratio of external to internal hours is roughly 6 times. For switchers, we find almost no scaling up.

The impact that investment has on growth depends importantly on the revenue share for the self-created intangibles, κ . In Figure 3.21, we report the predicted growth differential estimates as we vary ϕ and thus the revenue share. For the simulations, we hold all other parameters and prices fixed and thus find similar estimates for the incomes of our two groups at age 40, despite the fact that the life-cycle growth

patterns differ. As Panel A of Figure 3.21 shows, the choice of ϕ can have a large effect on the differential growth between entrepreneurs that continue and those that exit. In the baseline parameterization, we set ϕ to 0.15. In the figure, we show growth differentials as we lower ϕ to 0.1 and even further to 0.05. The associated investments in each case are shown in Panel B of Figure 3.21. With a larger revenue share for intangibles, the owner is incentivized to invest and the growth in income slower and more persistent. When the share is lowered, investments decline more quickly and the growth in incomes occurs earlier. In this case, the owners rely more on external factors and scale up the business at an earlier age. How much they scale up and when depends on the specification of hours in production, for example, the share of owner time, ω , and its substitutability with external labor, ρ . If owner and employee time is highly substitutable, then the owner can create the intangible asset—say the list of clients—and the employees can work with them. Relatedly, the external resource requirement versus own time in intangible capital production governed by ϑ is relevant for time use early in the career. But varying these parameters does not change the overall message that our predictions for the growth differentials depends importantly on incorporating nontrivial firm-specific investments.

Overall, we find that the model does surprisingly well in generating growth differentials that are consistent with the young entrepreneurs in the IRS sample.

3.6 Conclusions

Much has been written about the nature of entrepreneurship, but our knowledge base is built up from analyses of very different samples of individuals, which in the whole provide a narrative reminiscent of the parable of the blind men and an elephant. Each man learns about the elephant by touching only one part of the body, drawing conclusions that the elephant is like a wall, snake, spear, tree, fan, or rope depending on what they had touched. Analogously, the literature on entrepreneurship has an array of narratives, describing the typical business owner in many possible ways: as a gig worker seeking flexible arrangements; a misfit avoiding unemployment spells; an inventor seeking venture capital; a wealthy individual with no financing needs; a tax dodger. To provide a more complete picture of the nature of entrepreneurship, we used U.S. administrative tax data to assemble a novel longitudinal database of business owners—one that is suitable for analyzing patterns of income growth and determinants of entrepreneurial choice for a large population of self-employed.

Critical to the analysis was our notion of employment attachment, whether individuals were in the same employment status—say, self-, paid-, or non-employment—for most of the sample or were mostly switching. Comparisons of income and growth profiles for the attached self- and paid-employed revealed a striking contrast: the average income growth profiles of the self-employed are much higher and more hump-shaped than those for paid-employed peers with the same characteristics. Comparisons of

income changes for these groups revealed that dispersion in incomes declines over the life cycle regardless of employment status. Analysis of the switching into and out of self-employment also yielded new insights relative to earlier work. We found that individuals entering self-employment have higher past wage income and lower past asset income than peers not entering, which is contrary to earlier findings based on survey data.

We hope and expect that the empirical results of the paper will motivate new theories of entrepreneurship, which can be used to provide better tools for tax administrators and policymakers. In our view, critical inputs will include firm-specific investments, incomplete information about entrepreneurial productivity, and other mechanisms that imply a slow adjustment to the optimal size of operation.

3.7 Tables

Table 3.1: Main Sample Summary Statistics

Statistic	Total Sample	Attached		Almost		Mostly Switching	Any NE
		PE	SE	PE	SE		
Individuals (Mil.)	80.0	41.7	2.0	0.5	0.3	2.4	33.0
Shares (%)							
Counts	100	52.2	2.4	0.7	0.3	3.0	41.3
Total income	100	66.5	7.6	1.2	1.0	6.1	17.6
PE income	100	75.9	0.9	1.3	0.4	4.3	17.2
SE income	100	4.2	51.5	0.8	5.3	18.5	19.7
Incomes (2012\$, Thous.)							
Mean, PE plus SE income	49.0	62.5	152.1	89.5	146.9	98.7	20.8
Percentiles, 10 th	6.0	22.7	17.0	21.5	21.1	17.0	2.1
25 th	15.7	32.3	30.8	32.4	38.3	28.0	6.3
50 th	32.9	47.5	65.7	52.2	72.8	50.6	13.7
75 th	56.9	70.5	151.1	89.3	146.3	99.0	25.4
90 th	92.4	105.6	330.9	157.1	303.3	197.0	42.0
Mean, PE income	42.6	61.9	16.2	81.8	45.9	59.4	17.8
Percentiles, 10 th	3.3	22.6	0.0	20.4	4.5	10.3	0.8
25 th	12.8	32.3	0.0	30.8	8.2	17.7	4.4
50 th	30.3	47.4	2.0	48.6	17.5	31.7	11.8
75 th	53.4	70.1	10.4	80.7	41.4	58.7	23.1
90 th	84.4	104.6	32.6	139.6	91.0	108.9	38.1
Mean, SE income	6.5	0.5	135.9	7.7	101.0	39.3	3.1
Percentiles, 10 th	-0.0	-0.1	14.5	-4.5	10.0	0.4	0.0
25 th	0.0	0.0	26.5	0.1	20.4	5.6	0.0
50 th	0.0	0.0	57.3	3.1	45.7	14.3	0.0
75 th	0.4	0.0	134.2	8.8	101.3	37.5	0.7
90 th	6.8	0.8	297.8	21.0	218.0	90.0	5.9
Education and skills (%)							
College-educated	52.8	60.9	63.4	70.3	74.7	67.3	40.5
Cognitive	52.4	58.2	60.8	65.3	66.1	63.6	43.6
Interpersonal	58.6	66.5	61.2	69.6	71.4	66.4	47.7
Manual	37.6	35.3	35.3	35.0	28.0	34.3	41.0
Primary industry shares (%)							
Agriculture	1.1	0.9	1.6	1.3	1.6	1.2	1.3
Mining	0.4	0.5	0.5	0.7	0.5	0.6	0.3
Utilities	0.2	0.3	0.1	0.2	0.1	0.1	0.1

See notes at end of table.

Table 3.1: Main Sample Summary Statistics (Continued)

Statistic	Total Sample	Attached		Almost		Mostly Switching	Any NE
		PE	SE	PE	SE		
Primary industry shares (%)							
Construction	7.2	5.3	17.2	10.3	15.0	12.7	8.6
Manufacturing	11.5	14.3	3.6	10.5	4.7	6.3	8.9
Wholesale trade	3.6	4.0	3.4	4.2	4.0	3.8	3.2
Retail trade	8.9	8.1	7.9	8.4	9.5	8.4	10.0
Transportation	3.4	3.1	4.9	4.8	4.1	5.4	3.5
Information	1.6	1.8	0.9	2.5	1.3	1.6	1.4
Finance	3.3	3.7	4.5	4.3	4.4	4.2	2.6
Real estate	2.3	1.8	4.7	3.4	4.8	4.0	2.7
Professional services	9.1	8.8	16.0	15.3	17.2	16.6	8.2
Management	0.7	0.9	0.1	0.6	0.2	0.3	0.4
Administration	4.9	3.8	4.4	4.6	3.9	4.4	6.5
Education	0.6	0.5	0.4	0.6	0.5	0.7	0.6
Health care	6.4	5.5	10.6	6.4	10.3	9.3	7.1
Arts	1.0	0.8	1.8	1.5	2.1	1.6	1.2
Accommodation	4.4	3.3	3.4	4.0	4.8	4.2	6.0
Other services	3.8	2.1	11.1	4.2	9.0	7.7	5.3
Other NAICS	13.7	17.4	0.6	11.3	0.6	5.0	10.6
Missing NAICS	11.8	13.3	2.2	0.9	1.5	1.9	11.5
Demographics							
Male (%)	50.2	52.8	82.3	74.6	79.9	74.6	42.5
Mostly married (%)	63.8	68.1	78.7	72.4	81.8	73.7	56.4
Has children (%)	82.1	82.6	85.0	86.2	87.1	85.9	81.0
Number of children (mean)	2.2	2.2	2.2	2.5	2.3	2.4	2.3
Birth year (median)	1963	1963	1960	1964	1961	1963	1964
Other Incomes (2012\$, Thous.)							
Spousal wages, Mean	32.8	30.9	25.9	25.5	29.9	27.7	36.1
Median	17.9	20.7	8.6	10.6	13.4	11.3	15.7
Asset income, Mean	8.7	4.6	57.7	17.4	62.8	25.9	9.2
Median	0.0	0.1	1.9	0.1	2.6	0.3	0.0
UI income, Mean	0.4	0.3	0.1	0.5	0.1	0.3	0.5
Median	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Table 3.2: Largest Contributors to Attached Employee Growth Gap

Cumulative Share	Distinguishing Characteristics			
	NAICS	Cohort	Cognitive	Manual
9.6	62	1960s	✓	
18.8	54	1960s		
23.5	52	1960s		
27.7	62	1970s	✓	
31.5	54	1960s	✓	
34.3	44	1960s	✓	
36.9	62	1970s	✓	✓
39.5	54	1970s		
42.0	52	1970s		
44.6	62	1960s	✓	✓
47.1	23	1960s	✓	
49.4	23	1970s	✓	
51.7	54	1970s	✓	

Notes: The sample underlying this table includes all attached self- and paid-employed subgroups. Shares of the contribution to the self- and paid-employment growth gap—which is displayed in Panel B of Figure 3.2—are computed for ages between 30 and 39 and then cumulated. Results are reported for the top groups contributing at least 50 percent to the gap. All top contributors are male, married, college-educated, interpersonally-skilled, and have kids.

Table 3.3: Transition Probabilities for Incomes in Levels

		Attached Self-Employed										
		Income, Period t										
		<0	0- 7	7- 14	14- 27	27- 54	54- 106	106- 208	208- 410	410- 806	806- 1587	>1587
Income, Period $t - 1$	<0	49	2	5	8	10	10	7	4	2	1	1
	0- 7	6	19	33	22	12	5	2	1			
	7- 14	2	5	49	31	9	3	1				
	14- 27	2	1	13	57	22	4	1				
	27- 54	2	1	3	16	57	19	3	1			
	54- 106	2		1	3	17	58	16	2			
	106- 208	2			1	3	19	59	14	1		
	208- 410	2				1	3	19	60	12	1	
	410- 806	2				1	1	4	20	59	11	1
	806- 1587	3					1	2	5	22	55	12
	>1587	4					1	1	2	5	17	69
Distribution:		3	1	6	15	20	20	16	10	5	2	1

		Attached Paid-Employed										
		Income, Period t										
		<0	0- 7	7- 14	14- 27	27- 54	54- 106	106- 208	208- 410	410- 806	806- 1587	>1587
Income, Period $t - 1$	<0	29	7	11	19	20	10	3	1			
	0- 7	1	19	33	31	13	2					
	7- 14		5	46	36	11	1					
	14- 27		1	7	68	23	1					
	27- 54			1	7	81	10					
	54- 106				1	9	84	5				
	106- 208					1	14	78	6			
	208- 410						2	20	67	9	1	
	410- 806						1	5	24	59	9	1
	806- 1587						1	3	9	25	48	12
	>1587					1	1	3	7	11	21	56
Distribution:		0	1	3	14	39	32	9	2	0	0	0

Notes: The sample underlying these transition matrices includes all attached self- and paid-employed individuals. Incomes are in thousands of 2012 dollars. Element (i, j) of each matrix is the probability of having income in bin i in tax year $t - 1$ and income in bin j in tax year t . Elements with values below 0.5 percent are left blank. The stationary distribution is listed in the last lines of each matrix.

Table 3.4: Transition Probabilities for Residual Growth Rates

		Attached Self-Employed										
		% Change, Period t										
		<-628	-628- -185	-185- -54	-54- -16	-16- -5	-5- -5	-5- 16	16- 54	54- 185	185- 628	>628
% Change, Period $t - 1$	<-628	3	6	7	4	2	2	2	9	36	18	12
	-628- -185	2	8	11	5	2	2	2	8	30	25	6
	-185- -54	1	4	13	11	4	4	4	13	24	16	5
	-54- -16		1	10	21	10	9	10	21	15	3	
	-16- -5			7	22	15	16	14	18	7	1	
	-5- -5			6	22	18	17	14	16	6	1	
	-5- 16			7	24	17	16	13	16	6	1	
	16- 54		1	9	27	14	12	11	16	8	1	
	54- 185	1	2	16	27	10	8	8	14	10	3	1
	185- 628	1	4	23	22	7	6	6	13	12	5	1
	>628	2	5	28	20	6	5	5	11	11	5	1
Distribution:		0	2	10	22	12	11	10	16	11	4	1

		Attached Paid-Employed										
		% Change, Period t										
		<-628	-628- -185	-185- -54	-54- -16	-16- -5	-5- -5	-5- 16	16- 54	54- 185	185- 628	>628
% Change, Period $t - 1$	<-628	1	2	4	3	2	2	2	8	51	18	8
	-628- -185	1	3	5	4	1	1	2	8	39	33	5
	-185- -54			5	10	5	5	5	14	34	19	3
	-54- -16			4	16	15	17	15	21	12	1	
	-16- -5			2	10	34	23	18	10	2		
	-5- -5			1	8	45	20	18	7	1		
	-5- 16			2	11	35	23	19	9	1		
	16- 54			3	20	24	21	16	13	3		
	54- 185			6	23	15	19	14	15	6	1	
	185- 628			11	18	13	17	14	18	8	1	
	>628			17	17	11	14	12	18	9	2	
Distribution:		0	0	2	12	33	20	17	11	4	1	0

Notes: The sample underlying these transition matrices includes all attached self- and paid-employed individuals. Element (i, j) of each matrix is the probability of having the residual growth rate ($\Delta\epsilon$ divided by lagged income) in bin i in tax year $t - 1$ and the residual growth rate in bin j in tax year t . Elements with values below 0.5 percent are left blank. The stationary distribution is listed in the last lines of each matrix.

Table 3.5: Income Shares Held by Each Employment Group

Total Income							
Percentile Group	Total Sample	Attached PE	Attached SE	Almost PE	Almost SE	Mostly Switching	Any NE
$< 10^{th}$	-0.3	-0.2	-0.2	-0.0	-0.0	-0.1	0.2
$10^{th} - 25^{th}$	4.1	1.2	0.1	0.0	0.0	0.1	2.8
$25^{th} - 75^{th}$	35.5	24.2	0.9	0.3	0.1	1.2	8.7
$75^{th} - 90^{th}$	21.3	16.7	1.1	0.2	0.2	1.1	2.0
$> 90^{th}$	39	23.1	6.3	0.7	1.0	4.1	4.3

Paid-Employment Income							
Percentile Group	Total Sample	Attached PE	Attached SE	Almost PE	Almost SE	Mostly Switching	Any NE
$< 10^{th}$	1.1	0.1	0.0	0.0	0.0	0.0	0.9
$10^{th} - 25^{th}$	4.3	1.4	0.0	0.0	0.0	0.1	2.9
$25^{th} - 75^{th}$	38.5	28.0	0.1	0.3	0.0	1.0	9.0
$75^{th} - 90^{th}$	22.4	19.3	0.1	0.3	0.1	0.8	1.9
$> 90^{th}$	33.7	26.2	0.9	0.7	0.3	2.7	2.9

Self-Employment Income							
Percentile Group	Total Sample	Attached PE	Attached SE	Almost PE	Almost SE	Mostly Switching	Any NE
$< 10^{th}$	-8.7	-1.7	-1.7	-0.1	-0.2	-0.8	-4.2
$10^{th} - 25^{th}$	2.6	-0.0	0.4	-0.0	0.0	0.1	2.1
$25^{th} - 75^{th}$	16.7	0.7	5.8	0.1	0.6	2.9	6.6
$75^{th} - 90^{th}$	14.4	1.0	6.8	0.2	0.7	3.1	2.6
$> 90^{th}$	75.1	3.7	40.0	0.6	5.3	12.8	12.7

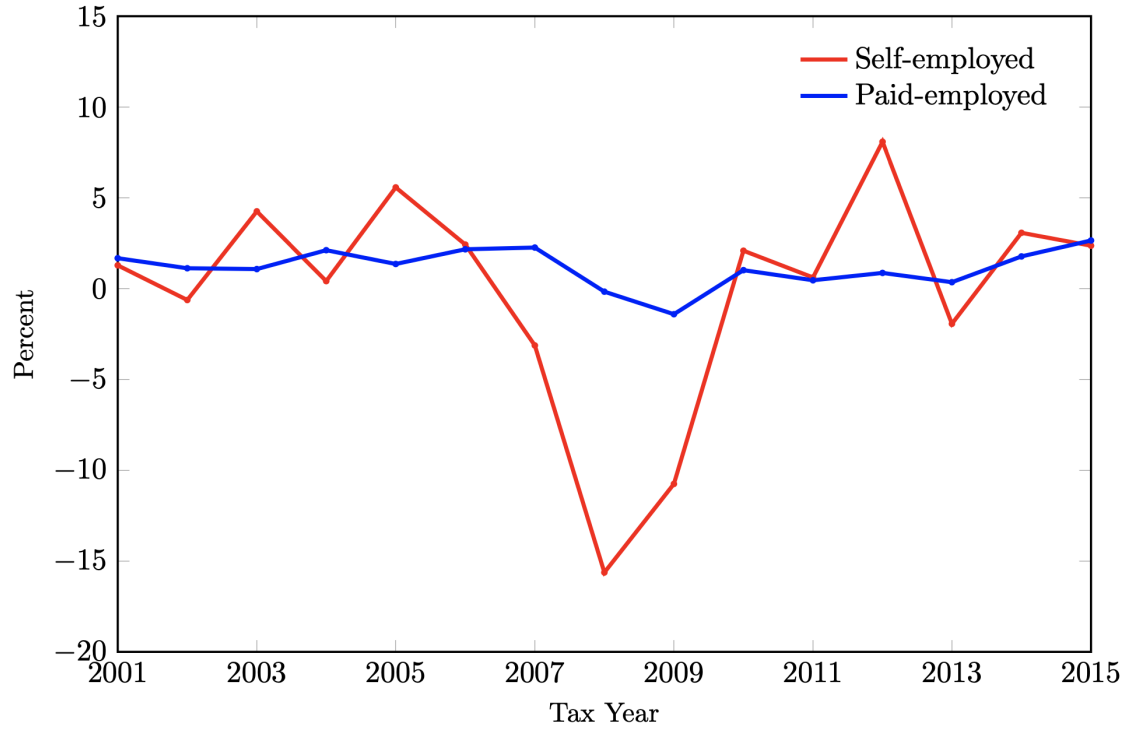
Notes: The sample underlying this table includes all individuals in the baseline sample summarized in the first column of Table 3.1. The income shares are total income attributed to the group listed in the column headings divided by the total in the ranked group.

Table 3.6: Model Parameters

Parameter	Expression	Value
Preferences		
Leisure weight	ψ	0.580
Intertemporal elasticity inverse	σ	1.500
Discount factor	β	0.960
Technologies		
Owner hours share, intangible production	ϑ	0.408
Hours substitution parameter, goods production	ρ	0.500
Intangible capital share, goods production	ϕ	0.150
Fixed asset share, goods production	α	0.300
Owner hours share, goods production	ω	0.425
Intangible capital depreciation	δ_{κ}	0.058
Fixed asset depreciation	δ_k	0.041
Prices		
Interest rate	r	0.041
Hired labor	w	1.000
Goods and services	p	1.500
Intangible capital	p_k	1.600
Entrepreneurial productivity		
Initial predicted mean	μ_0	0.000
Initial predicted variance	σ_0^2	0.005
Idiosyncratic shock variance	σ_{η}^2	0.004
Paid-employment shocks		
Persistence	ρ_{ϵ}	0.700
Standard deviation	σ_{ϵ}	0.100

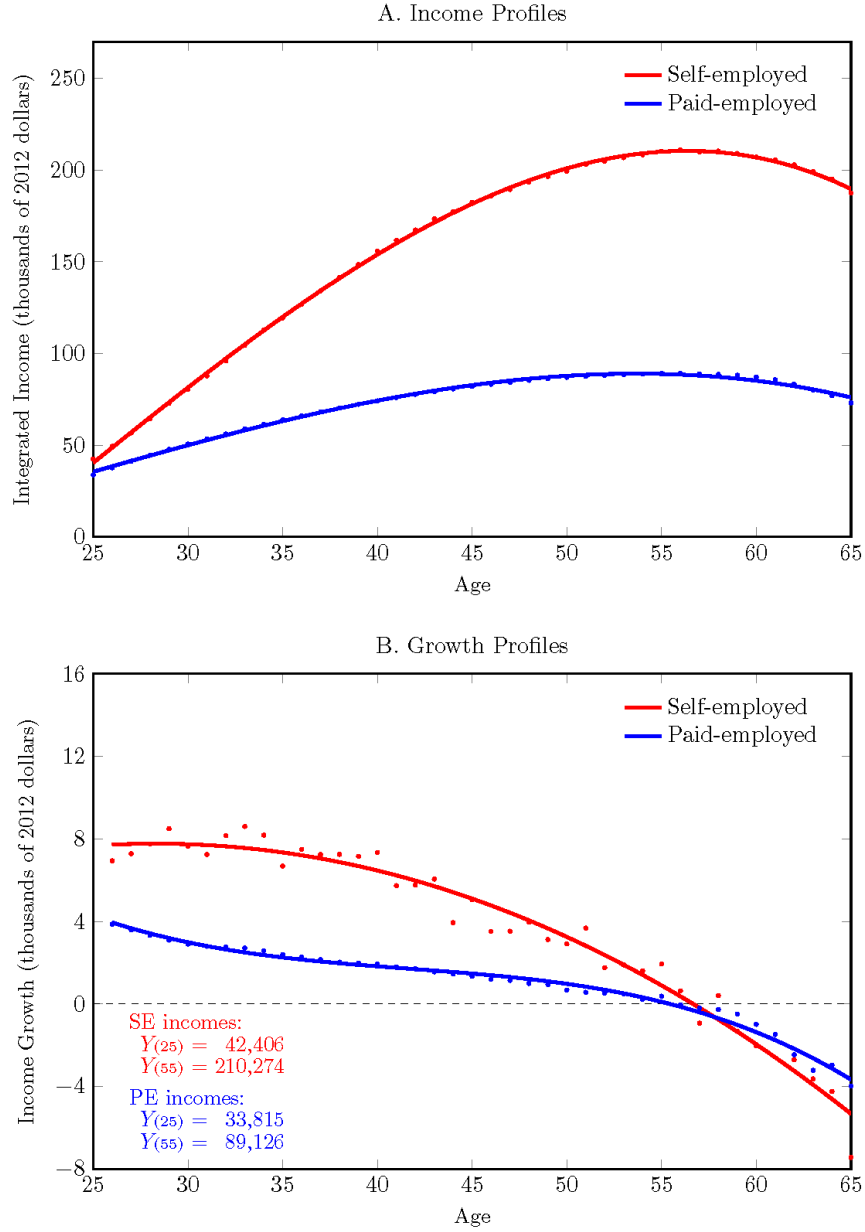
3.8 Figures

Figure 3.1: Estimated Time Effects Relative to Average Income



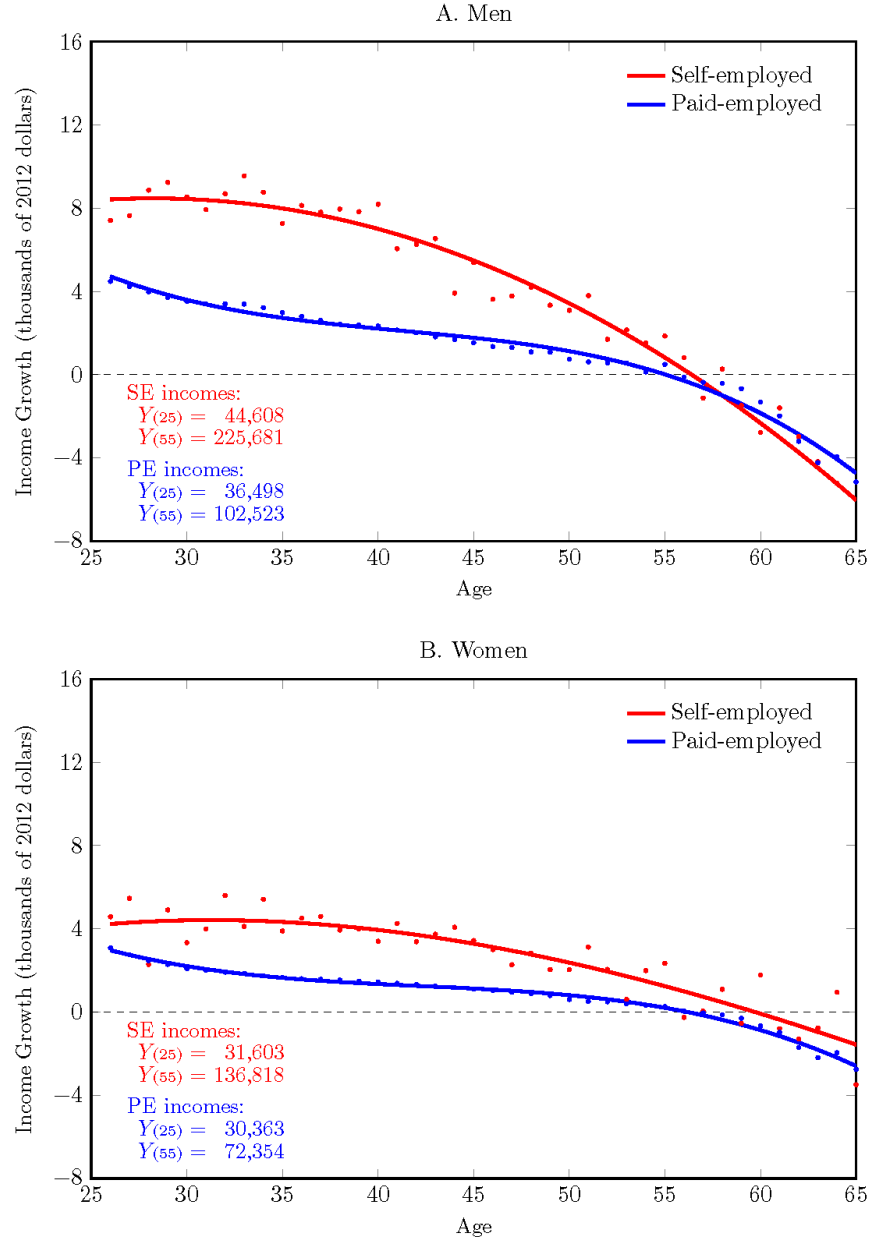
Note: The sample underlying these figures includes all attached self- and paid-employed subgroups. The figure reports weighted averages of the estimated time effects for groups g at time t , that is, $\Delta\beta_{g,t}$, which is divided by average income for group g in year t , $\bar{y}_{g,t}$. Weights are constructed from group counts.

Figure 3.2: Income and Growth Profiles



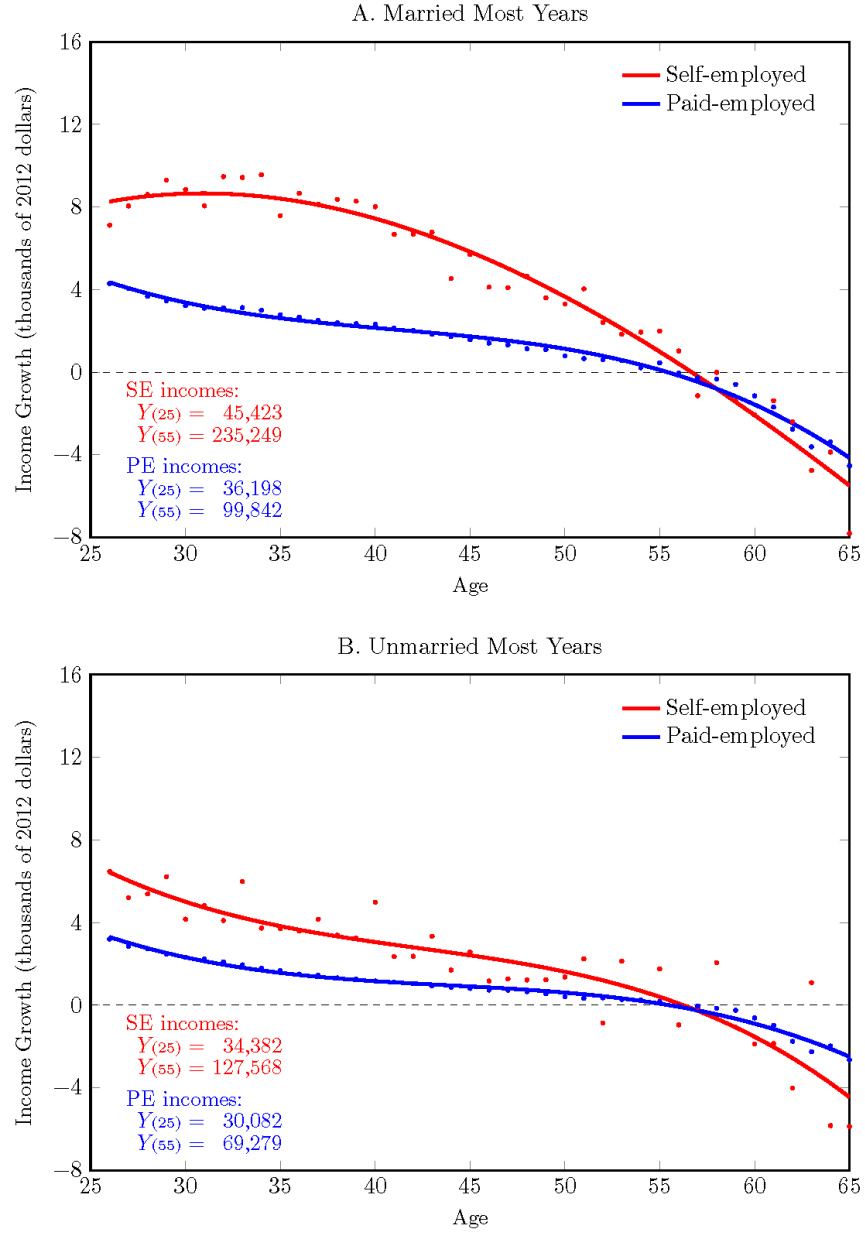
Notes: The sample underlying these figures includes all attached self- and paid-employed subgroups. Panel A shows the weighted averages of integrated incomes $Y_g(a)$ for subgroup g at age a . Panel B shows weighted averages of the associated growth by age, $Y_g(a) - Y_g(a-1)$. In both panels, weights are constructed from sample counts, $N_g^a / \sum_g N_g^a$.

Figure 3.3: Growth Profiles by Gender



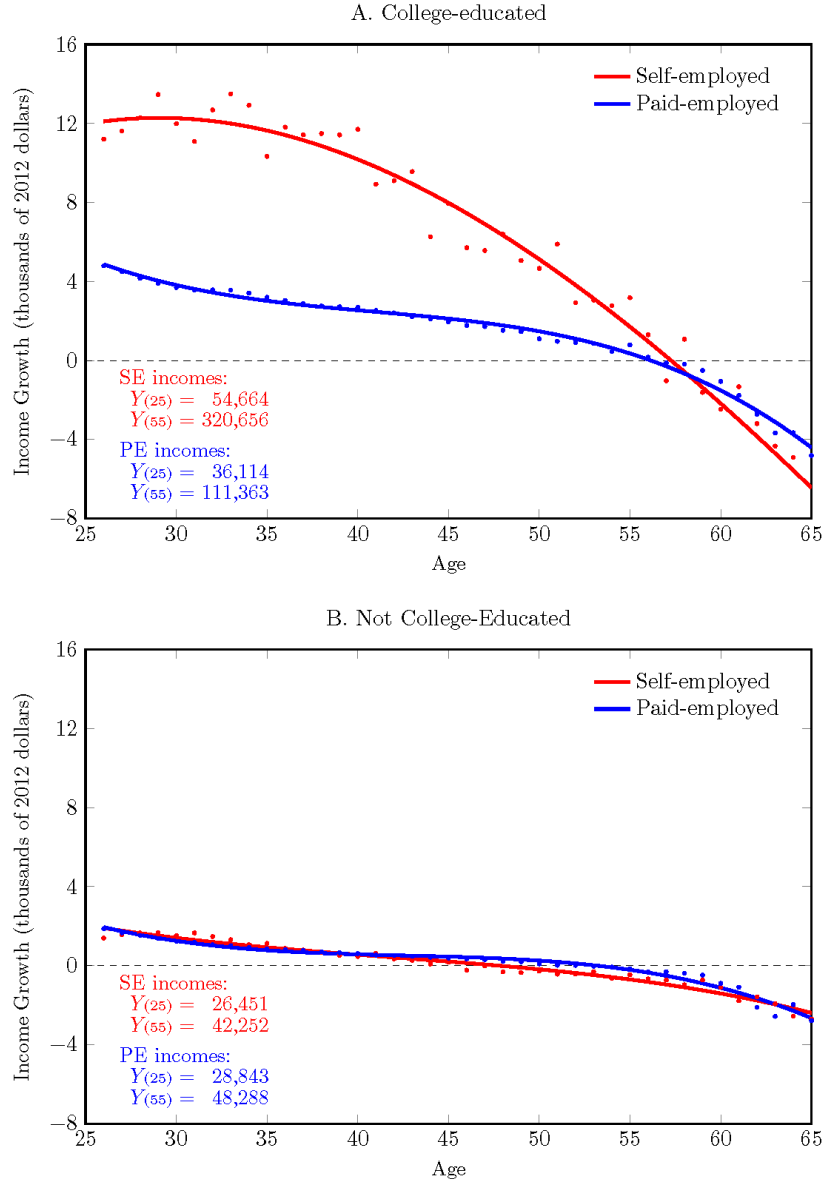
Notes: The sample underlying these figures includes all attached self- and paid-employed subgroups, which are further subdivided into those for men (Panel A) and women (Panel B). Both panels show weighted averages of subgroup g growth for each age a , $Y_g(a) - Y_g(a-1)$, with weights constructed from sample counts, $N_g^a / \sum_g N_g^a$.

Figure 3.4: Growth Profiles by Marital Status



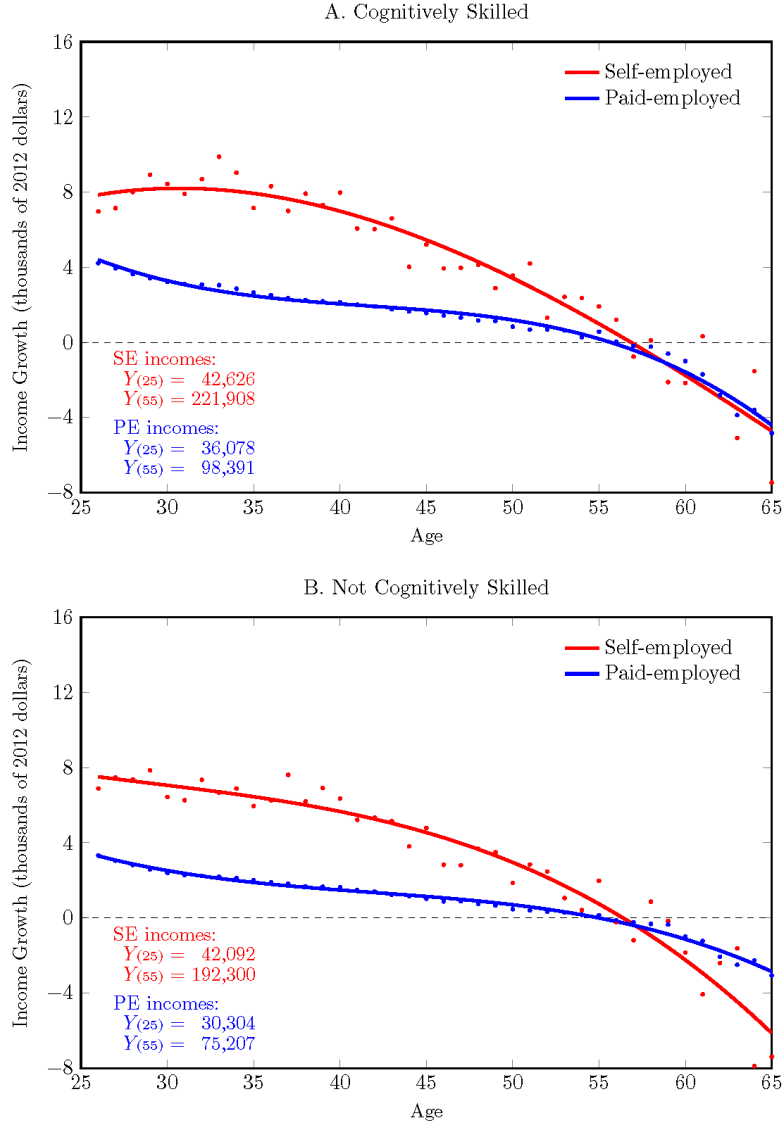
Notes: The sample underlying these figures includes all attached self- and paid-employed subgroups, which are further subdivided into those for the married most years (Panel A) and those that are not (Panel B). Both panels show weighted averages of subgroup g growth for each age a , $Y_g(a) - Y_g(a-1)$, with weights constructed from sample counts, $N_g^a / \sum_g N_g^a$.

Figure 3.5: Growth Profiles by Education



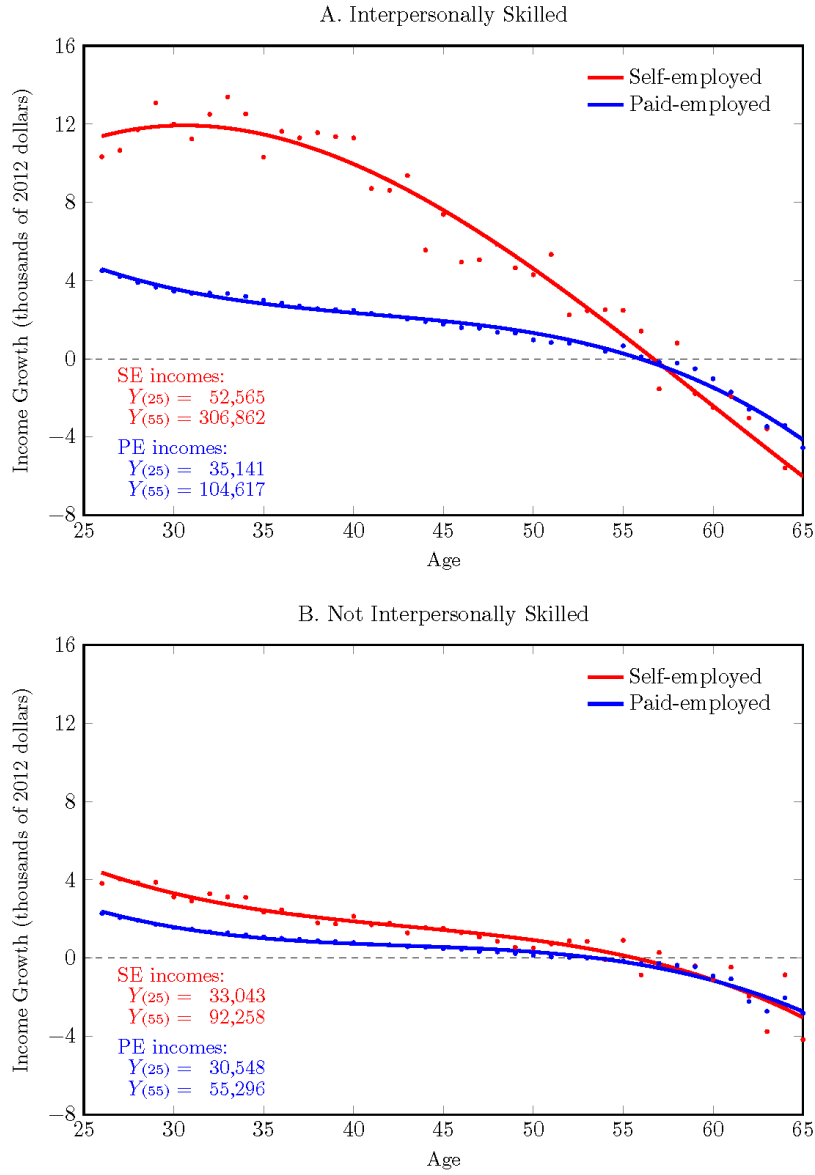
Notes: The sample underlying these figures includes all attached self- and paid-employed subgroups, which are further subdivided into those for classified as college-educated (Panel A) and those that are not (Panel B). Both panels show weighted averages of subgroup g growth for each age a , $Y_g(a) - Y_g(a-1)$, with weights constructed from sample counts, $N_g^a / \sum_g N_g^a$.

Figure 3.6: Growth Profiles by Cognitive Skill



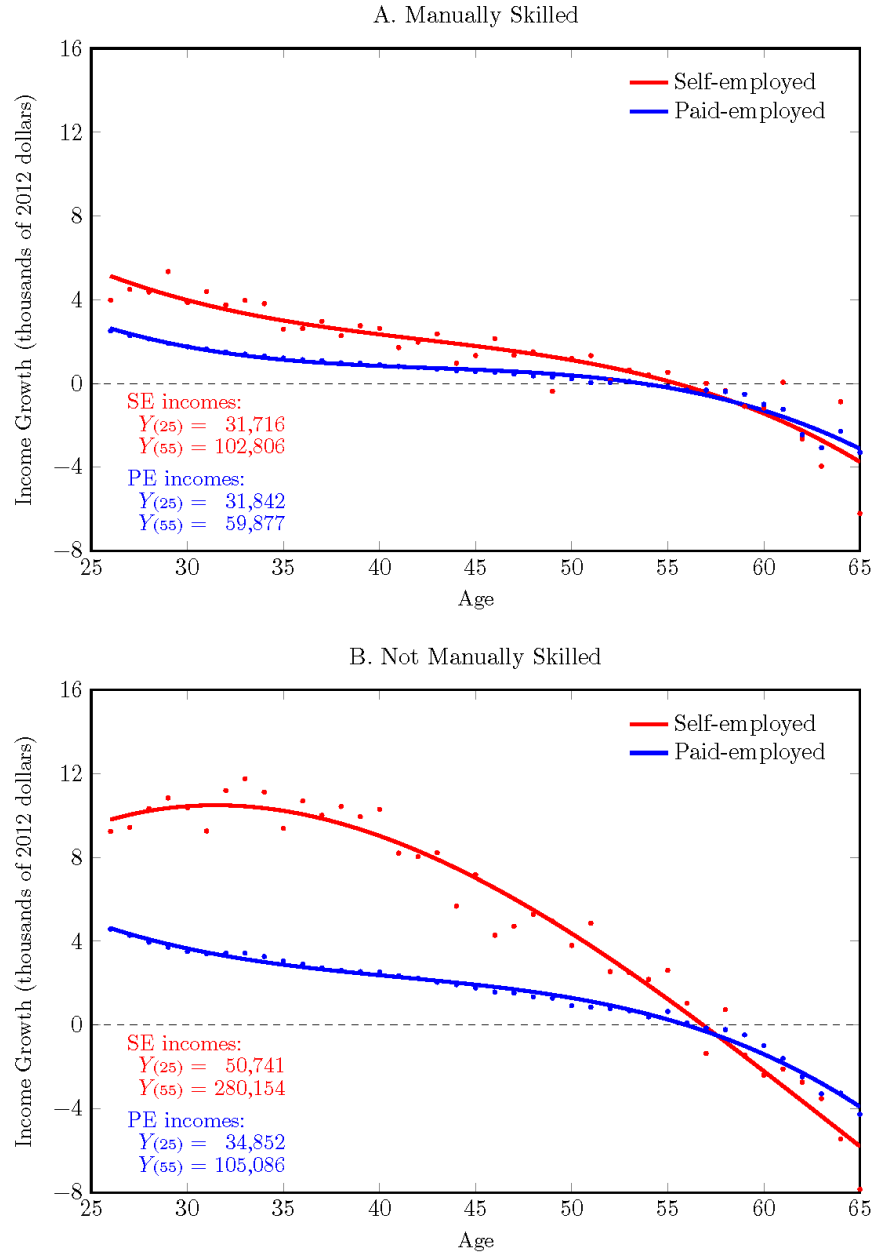
Notes: The sample underlying these figures includes all attached self- and paid-employed subgroups, which are further subdivided into those for classified as cognitively skilled (Panel A) and those that are not (Panel B). Both panels show weighted averages of subgroup g growth for each age a , $Y_g(a) - Y_g(a-1)$, with weights constructed from sample counts, $N_g^a / \sum_g N_g^a$.

Figure 3.7: Growth Profiles by Interpersonal Skill



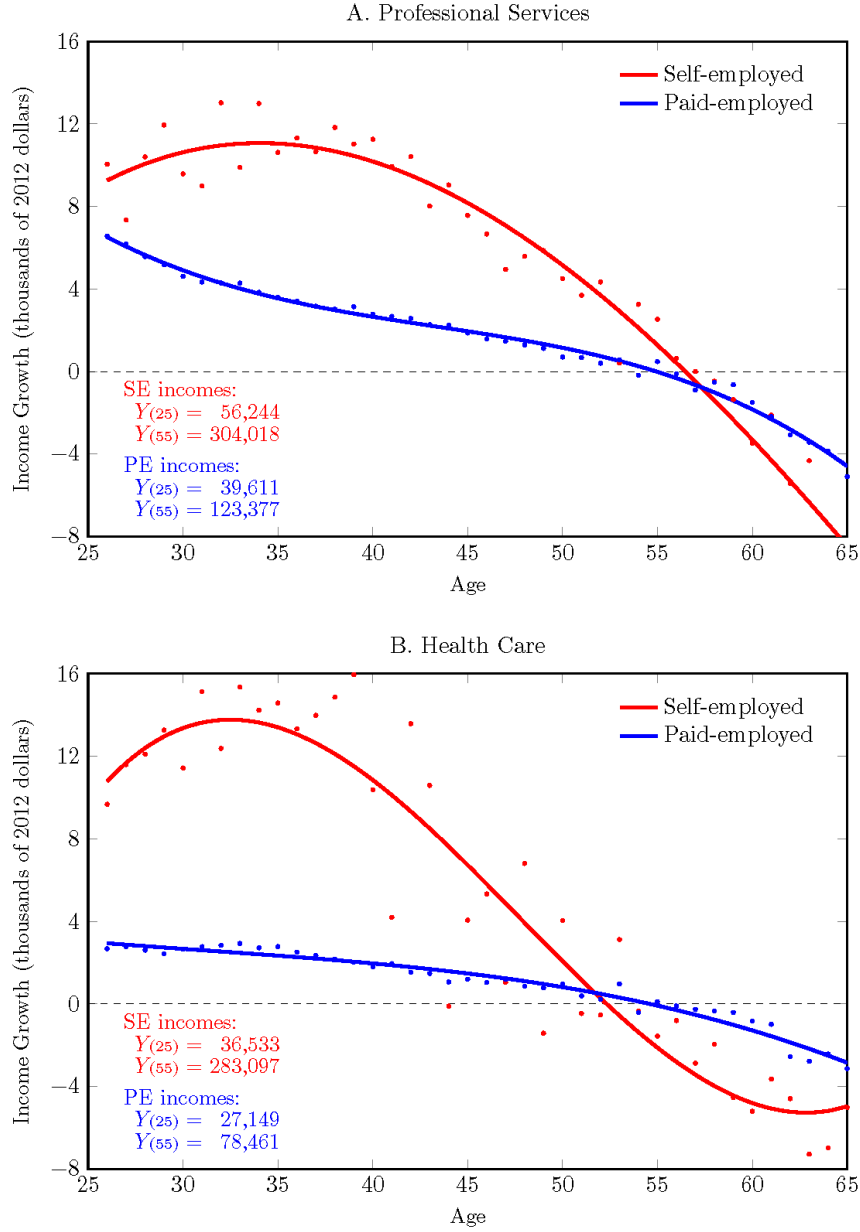
Notes: The sample underlying these figures includes all attached self- and paid-employed subgroups, which are further subdivided into those for classified as interpersonally skilled (Panel A) and those that are not (Panel B). Both panels show weighted averages of subgroup g growth for each age a , $Y_g(a) - Y_g(a-1)$, with weights constructed from sample counts, $N_g^a / \sum_g N_g^a$.

Figure 3.8: Growth Profiles by Manual Skill



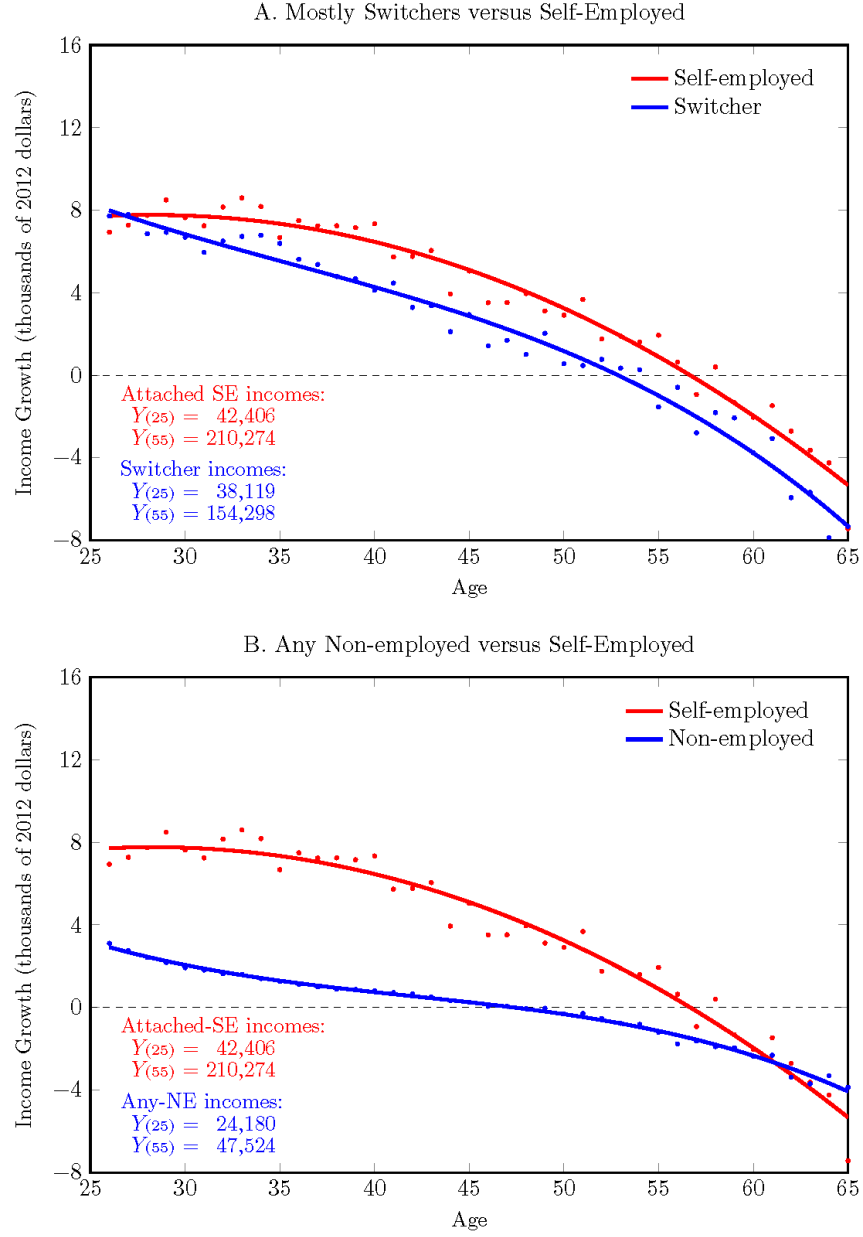
Notes: The sample underlying these figures includes all attached self- and paid-employed subgroups, which are further subdivided into those for classified as manually skilled (Panel A) and those that are not (Panel B). Both panels show weighted averages of subgroup g growth for each age a , $Y_g(a) - Y_g(a-1)$, with weights constructed from sample counts, $N_g^a / \sum_g N_g^a$.

Figure 3.9: Growth Profiles by Industry



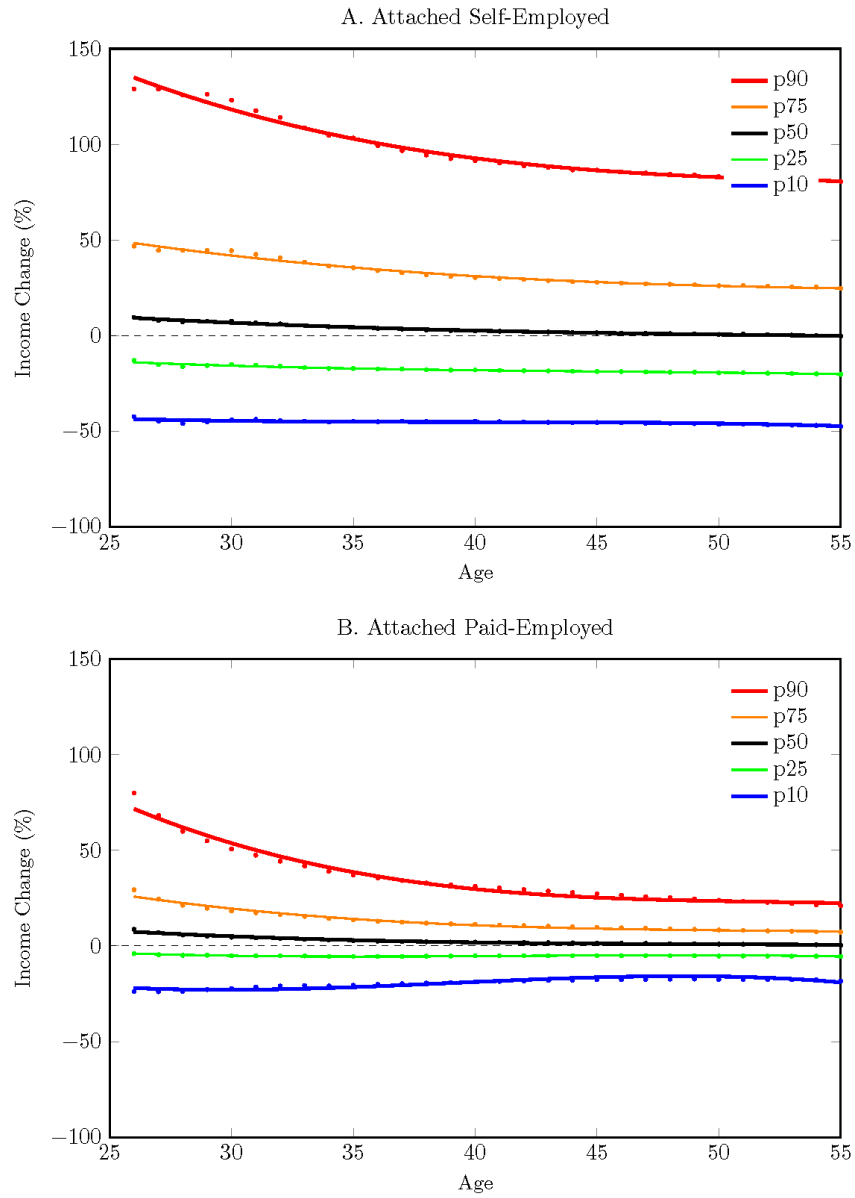
Notes: The sample underlying these figures includes individuals in the attached self- and paid-employed subgroups that work in professional services (Panel A) and health care (Panel B). Both panels show weighted averages of subgroup g growth for each age a , $Y_g(a) - Y_g(a-1)$, with weights constructed from sample counts, $N_g^a / \sum_g N_g^a$.

Figure 3.10: Growth Profiles of Mostly Switchers and Any Non-Employed



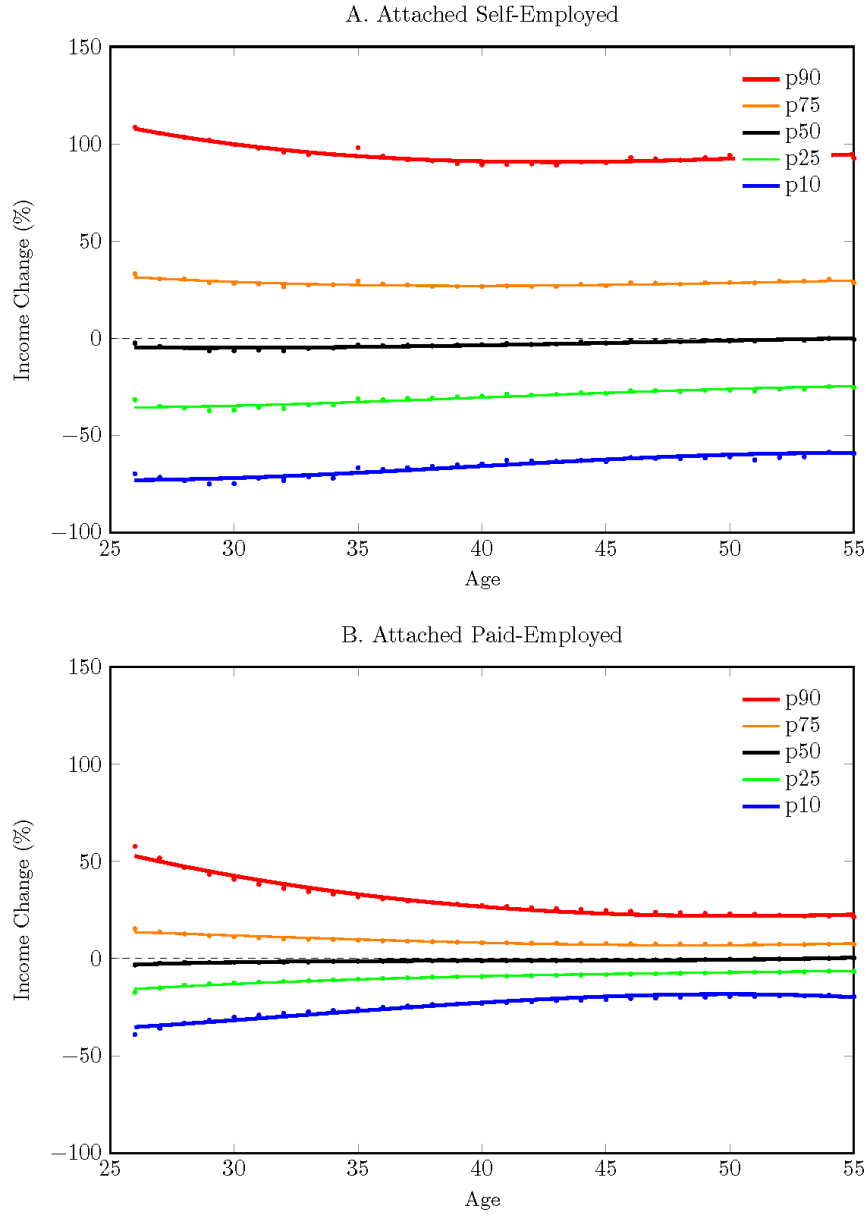
Notes: The sample underlying these figures includes individuals in the mostly-switchers and any-non-employment subgroups. Both panels show weighted averages of subgroup g growth for each age a , $Y_g(a) - Y_g(a-1)$, with weights constructed from sample counts, $N_g^a / \sum_g N_g^a$. Results are compared to the growth of the attached self-employed from Figure 3.2.

Figure 3.11: Age-Over-Age Growth in Incomes



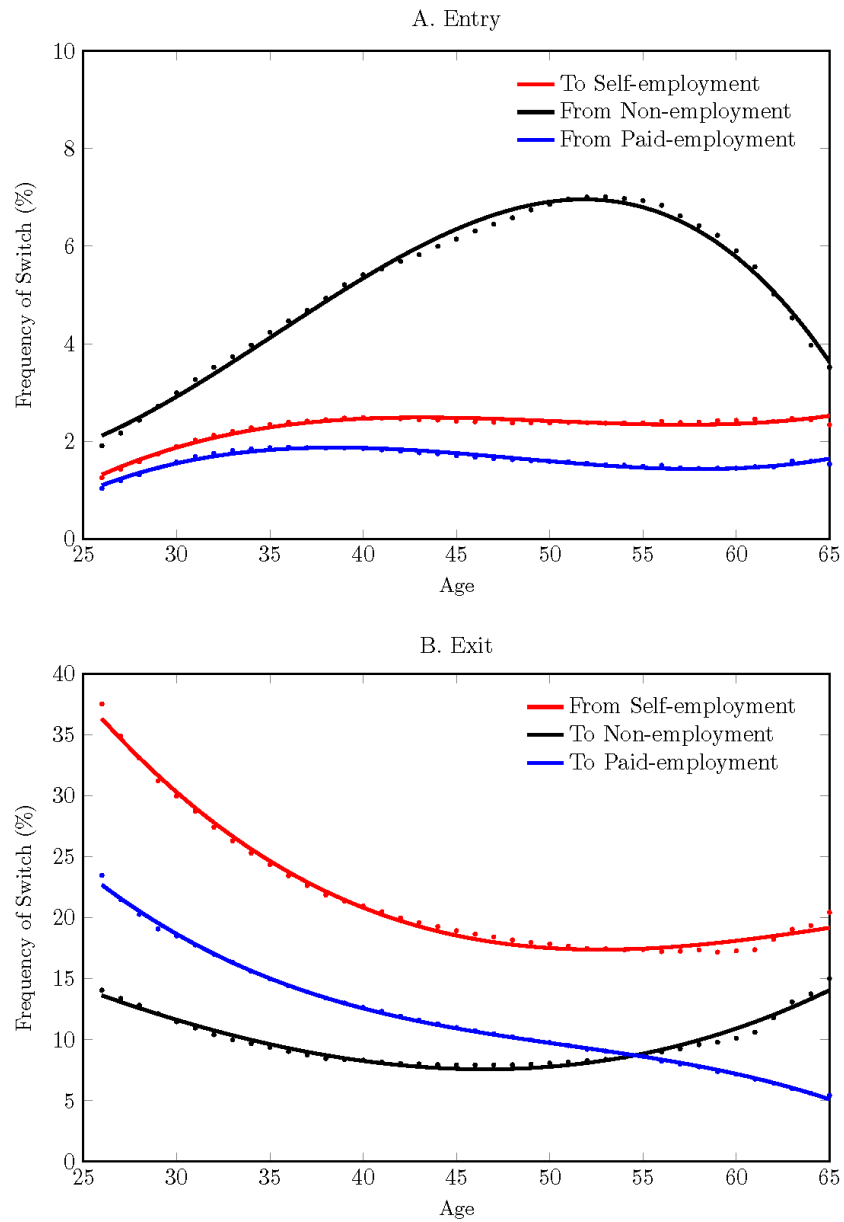
Notes: The sample underlying these figures includes all attached self- and paid-employed individuals. For each individual, we compute the age-over-age percentage change in income, $\Delta y_{ia}/|y_{i,a-1}|$, and plot selected percentiles of these changes for the self-employed (Panel A) and paid-employed (Panel B).

Figure 3.12: Age-Over-Age Growth in Regression Residuals



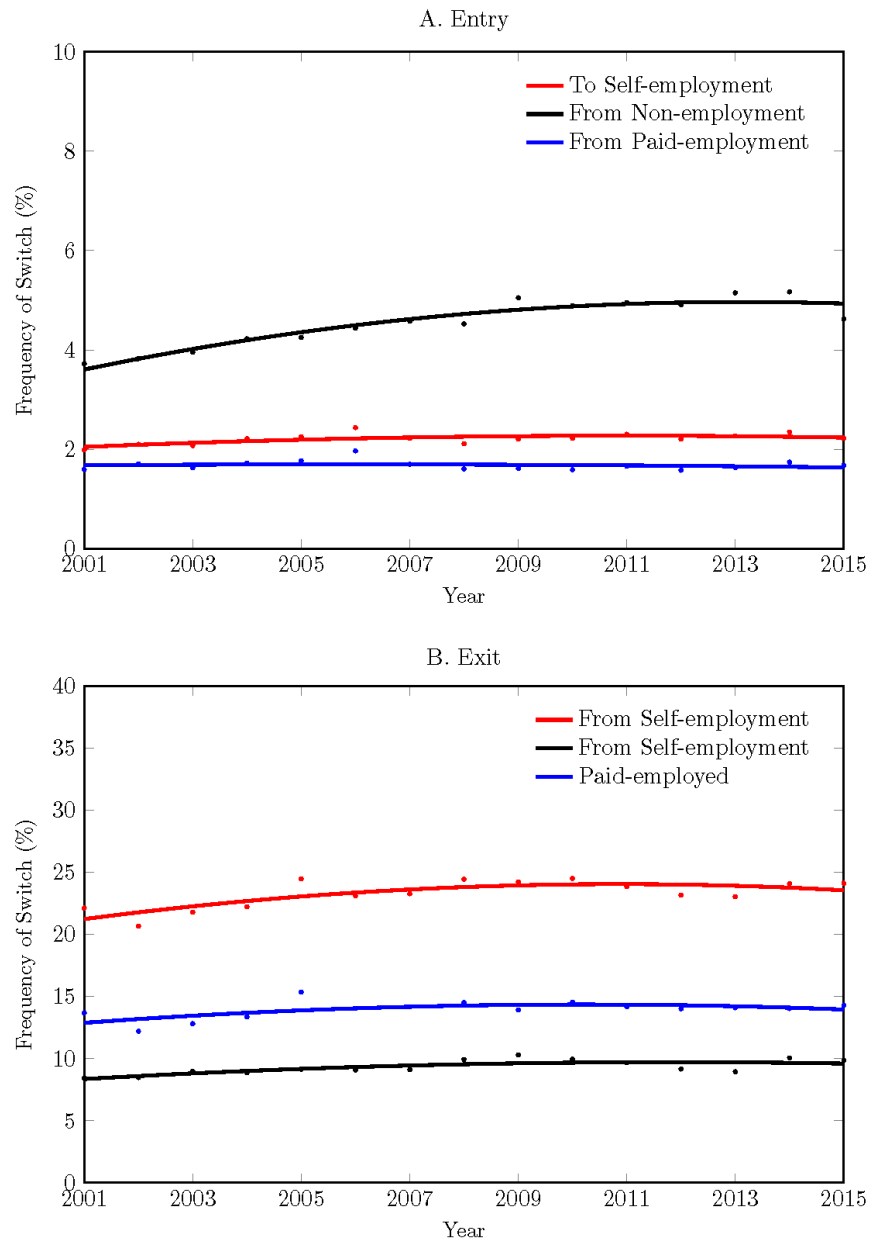
Notes: The sample underlying these figures includes all attached self- and paid-employed individuals. For each individual, we compute the age-over-age change in the residual normalized by previous year income, $\Delta\epsilon_{ia}/|y_{i,a-1}|$ and plot selected percentiles of these changes for the self-employed (Panel A) and paid-employed (Panel B).

Figure 3.13: Self-Employment Switching Rates By Age



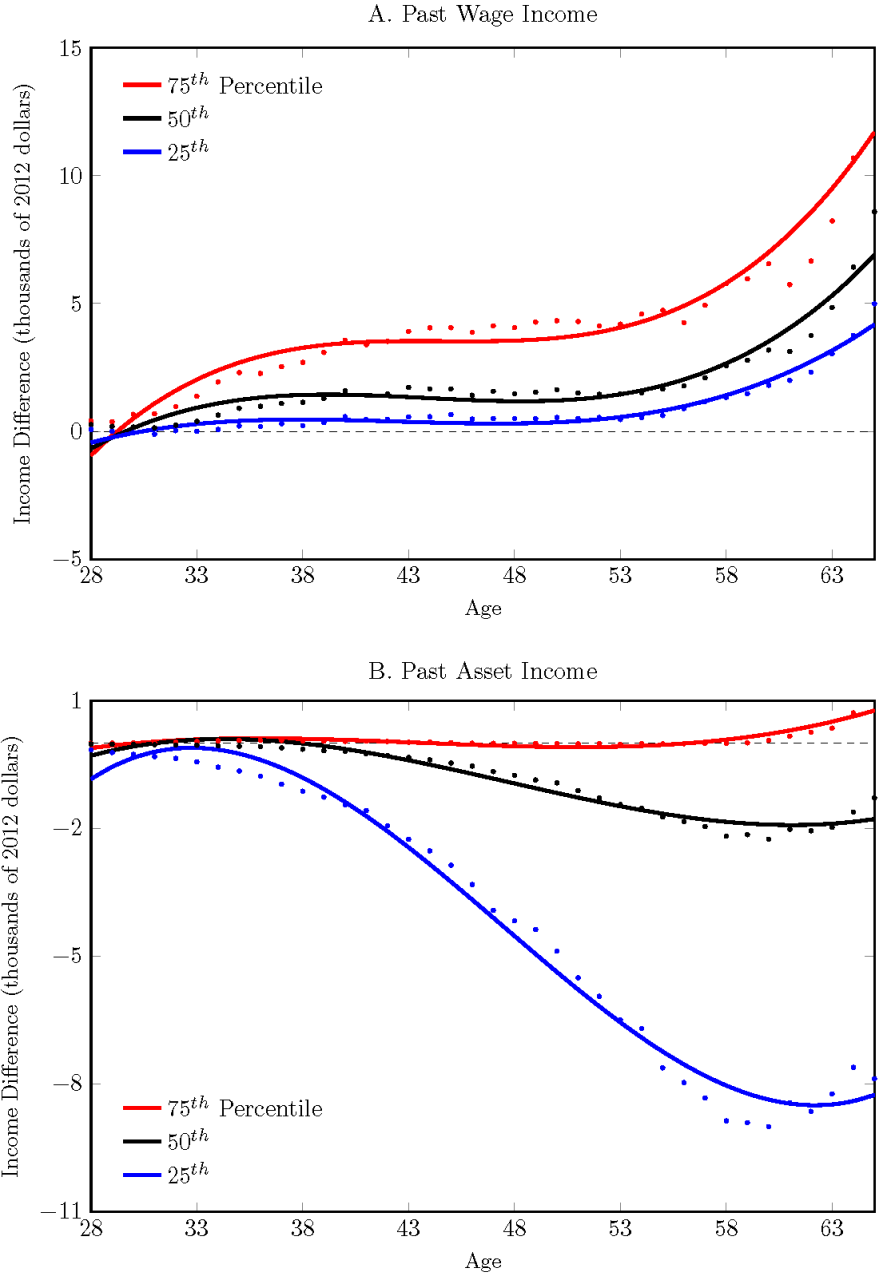
Notes: The sample underlying these figures includes all individuals in the Total Sample of Table 3.1. Entry rates into self-employment are shown in Panel A, for all non-self-employed and separately for the paid- and non-employed. Exit rates are shown in Panel B, with transitions to paid- and non-employed shown separately.

Figure 3.14: Self-Employment Switching Rates By Year



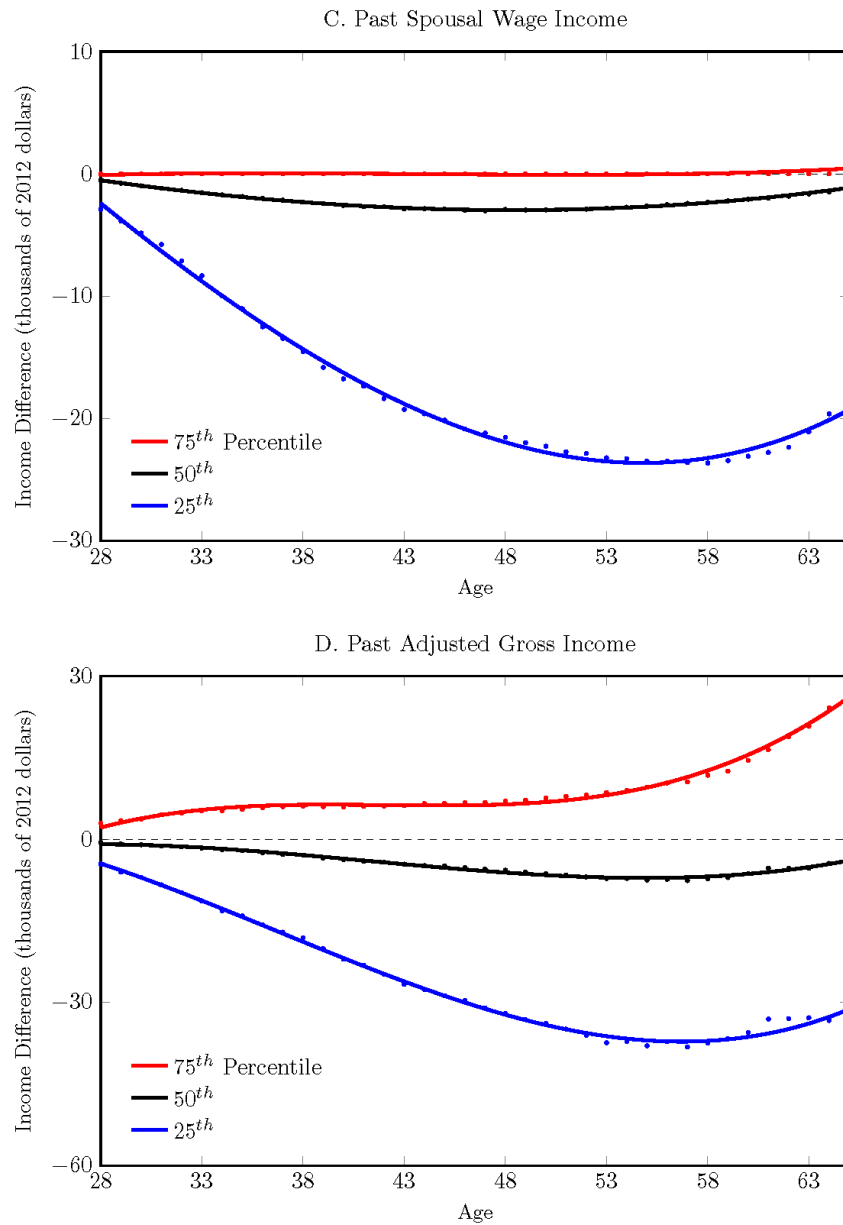
Notes: The sample underlying these figures includes all individuals in the Total Sample of Table 3.1. Entry rates into self-employment are shown in Panel A, for all non-self-employed and separately for the paid- and non-employed. Exit rates are shown in Panel B, with transitions to paid- and non-employed shown separately.

Figure 3.15: Differences in Past Incomes Between Current and Future Switchers



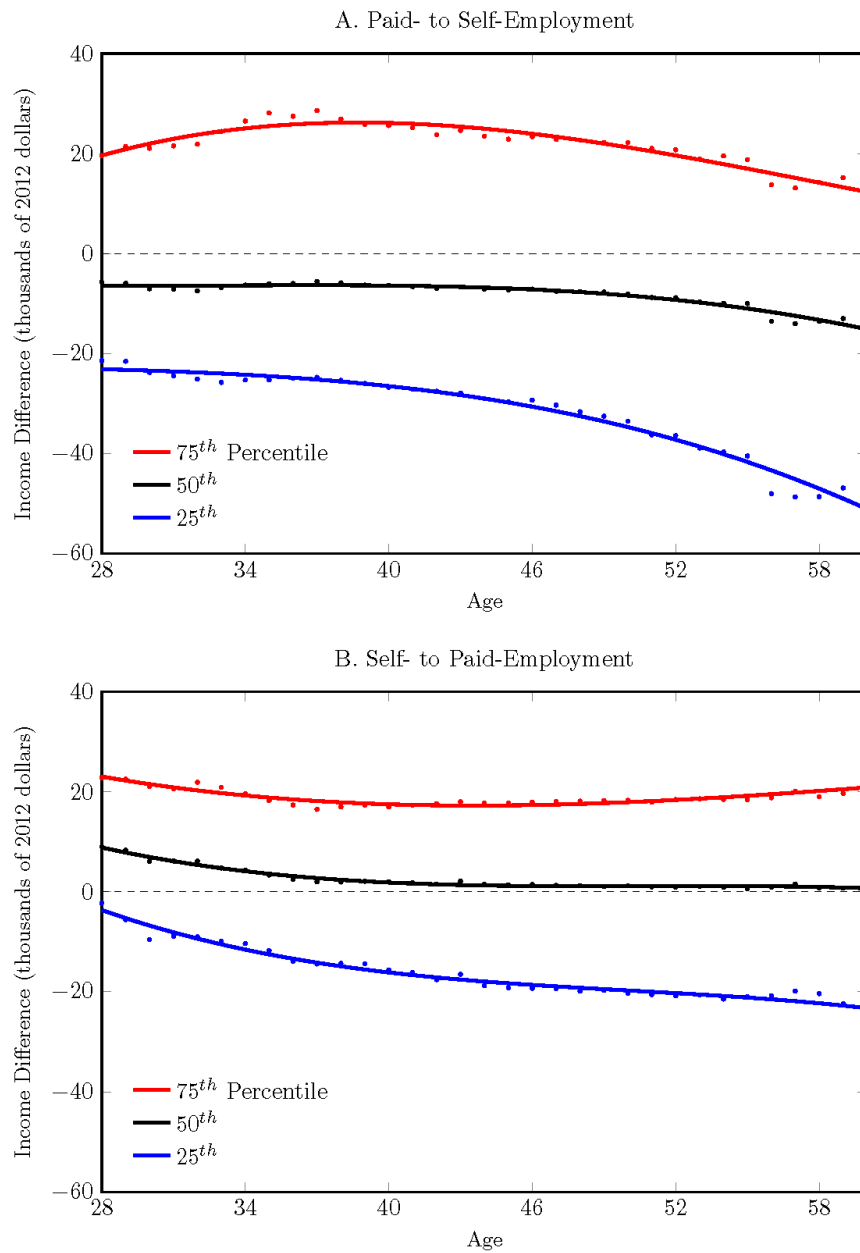
See notes at the end of the table.

Figure 3.15: Differences in Past Incomes Between Current and Future Switchers (Cont.)



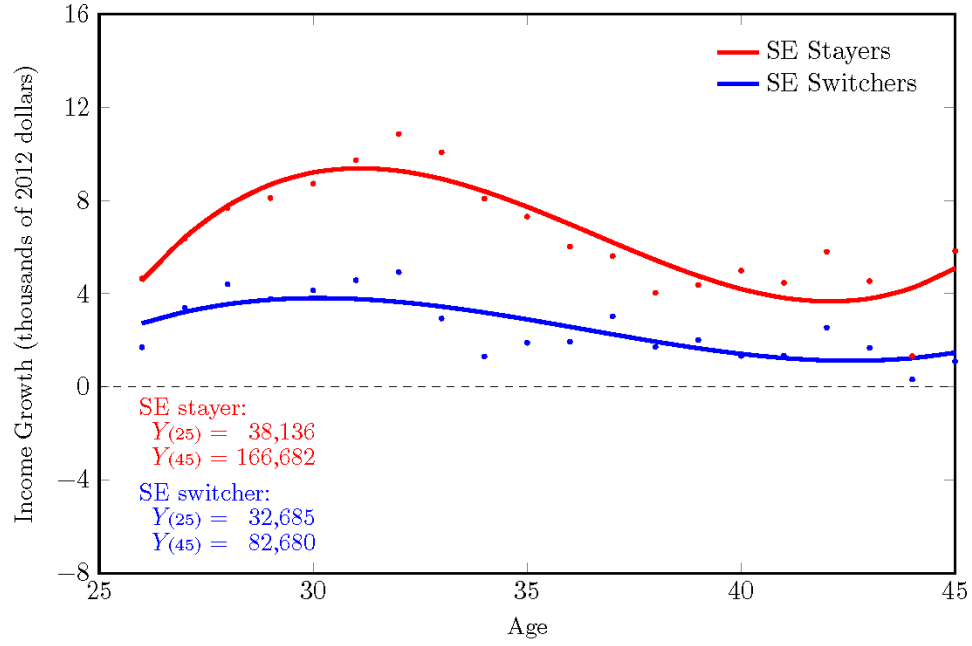
Notes: The sample underlying these figures includes individuals with at most one observed switch between paid- and self-employment. Each panel displays the interquartiles of differences in average past incomes at each age, that is, the average wage of the switcher less the average wage of peers that have similar characteristics but switch later.

Figure 3.16: Differences in Incomes Between Switchers and Non-switchers



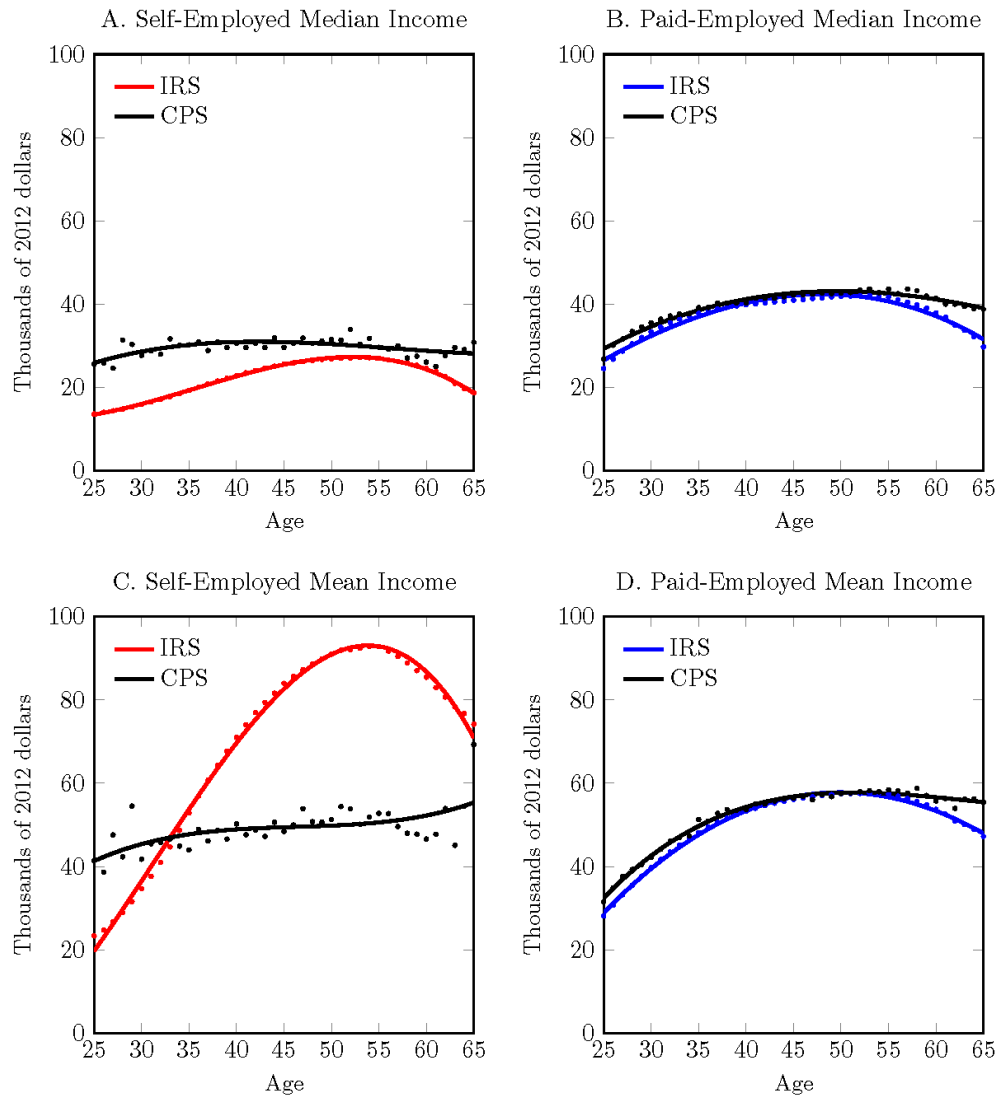
Notes: The sample underlying these figures includes individuals that switch at least once between paid- and self-employment. Panel A shows the income difference—averaged over 3 years before and after the switch—for those switching from paid- to self-employment. Panel B shows the income difference for this switching from self- to paid-employment.

Figure 3.17: Growth Differentials for Young Entrepreneurs



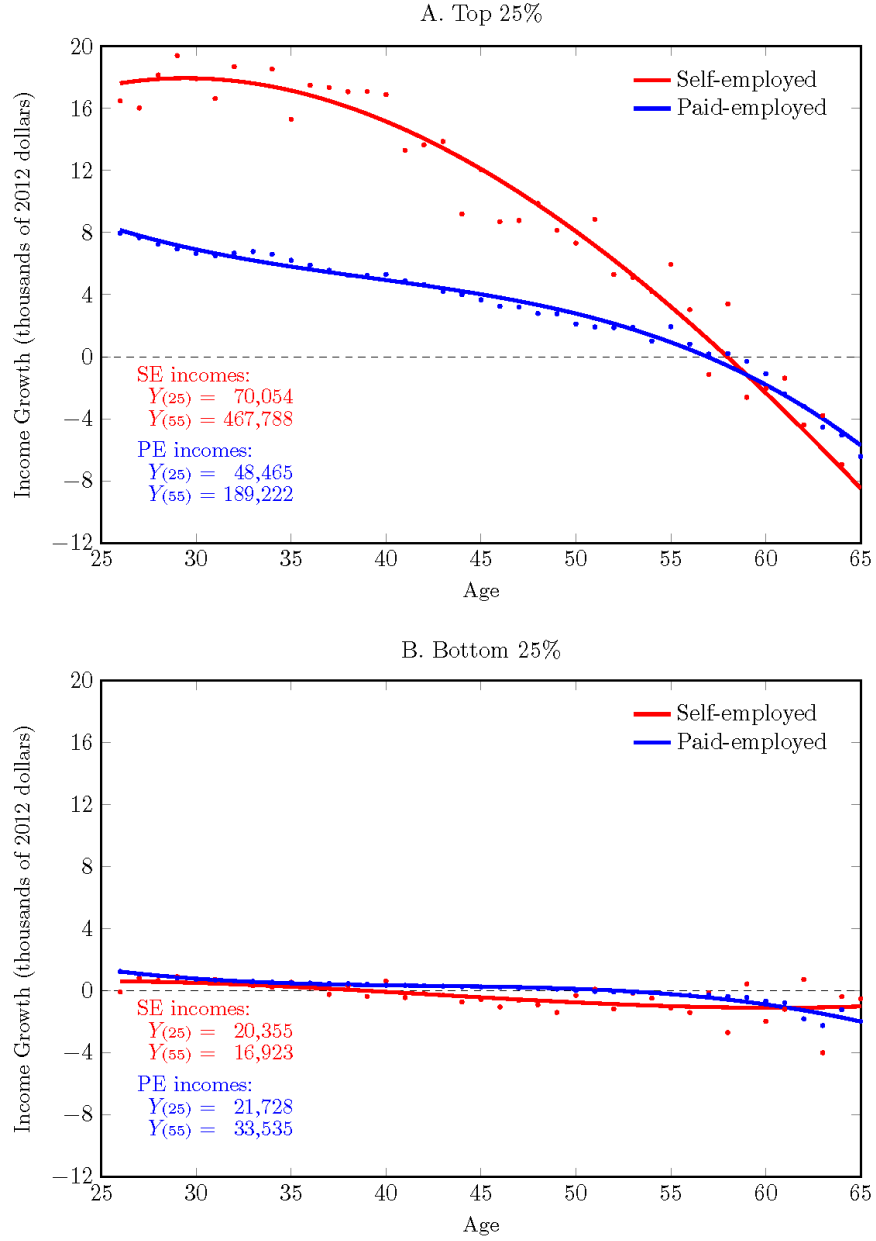
Notes: The sample underlying these figures includes all individuals born in the 1970–1975 cohorts with at least five years of self-employment experience before age 35. Growth profiles are plotted separately for those who, after age 35, remained in self-employment and those who switched to paid-employment. Growth in income for each age a , $Y_g(a) - Y_g(a-1)$, is computed for all subgroups g and averaged using count weights, $N_g^a / \sum_g N_g^a$.

Figure 3.18: Empirical Moments, IRS versus CPS



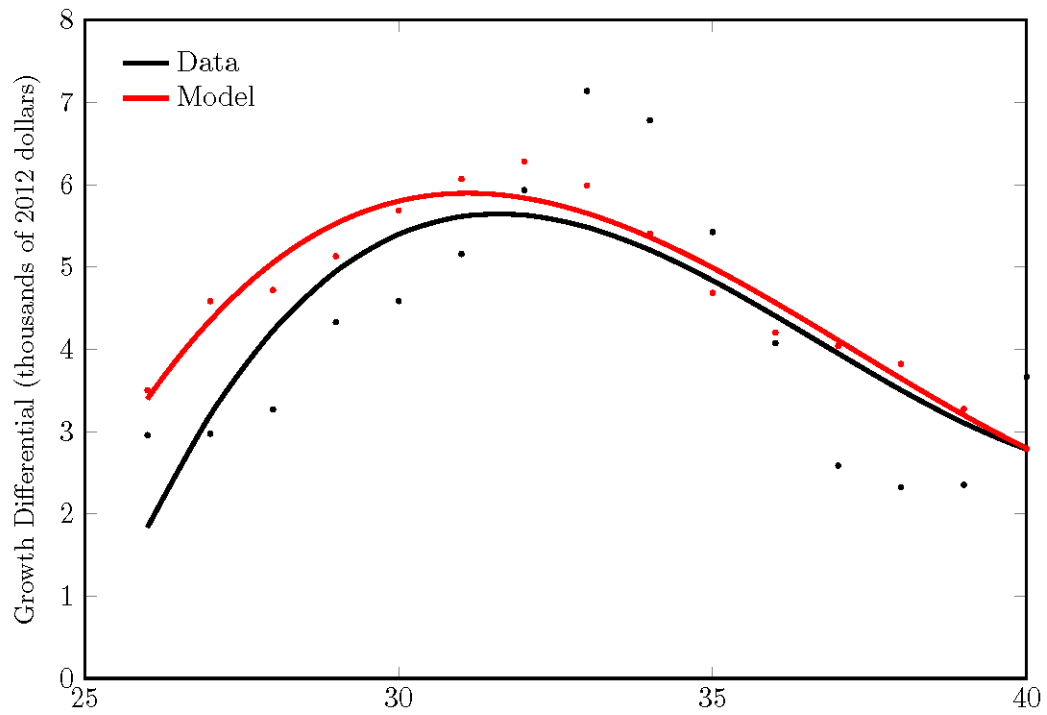
Notes: For both the IRS and CPS samples, individuals are assigned to self-employment in a particular year/age if the absolute value of income from business exceeds \$5,000 in 2012 dollars and the income from paid-employment. If these criteria are not met but income from non-business wages and salaries exceeds \$5,000 (in 2012 dollars), then they are assigned to paid-employment. The sample is not balanced and statistics are computed for each age.

Figure 3.19: Growth Profiles for Top and Bottom 25% Income Ranks



Notes: The sample underlying these figures includes individuals in the attached self- and paid-employed subgroups that are ranked by income into the top 25 percent (Panel A) or bottom 25 percent (Panel B). Both panels show weighted averages of subgroup g growth for each age a , $Y_g(a) - Y_g(a-1)$, with weights constructed from sample counts, $N_g^a / \sum_g N_g^a$.

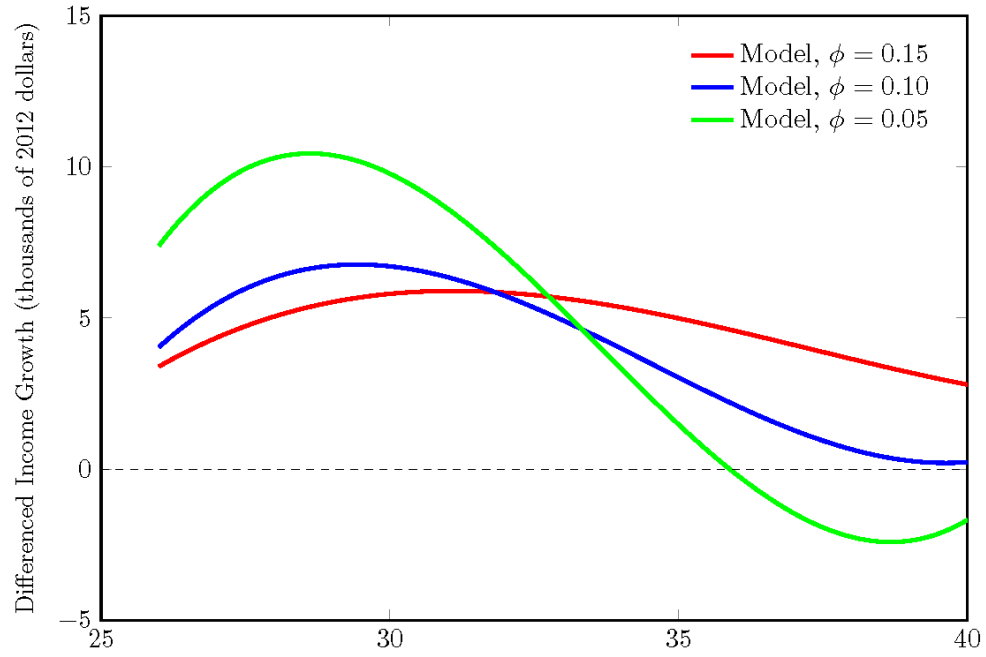
Figure 3.20: Growth Differentials for Young Entrepreneurs



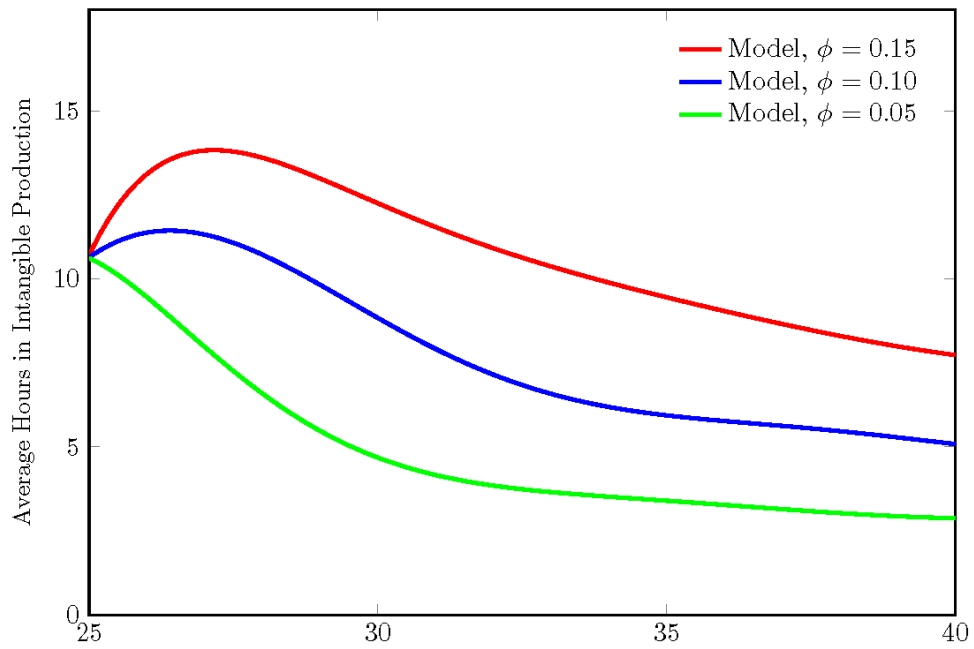
Note: See Figure 3.17 and Section 3.4.3 for description of data.

Figure 3.21: Model Predictions as Intangible Revenue Varied

A. Growth Differences



B. Owner Hours of Investment



References

- John M. Abowd and David Card. On the covariance structure of earnings and hours changes. *Econometrica*, 57(2):411–445, 1989.
- Daron Acemoglu, Ufuk Akcigit, Harun Alp, Nicholas Bloom, and William Kerr. Innovation, reallocation, and growth. *American Economic Review*, 108(11):3450–91, November 2018. doi: 10.1257/aer.20130470. URL <https://www.aeaweb.org/articles?id=10.1257/aer.20130470>.
- Philippe Aghion, Nick Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt. Competition and Innovation: an Inverted-U Relationship. *The Quarterly Journal of Economics*, 120(2):701–728, 05 2005.
- Ufuk Akcigit, Murat Alp Celik, and Jeremy Greenwood. Buy, keep, or sell: Economic growth and the market for ideas. *Econometrica*, 84(3):943–984, 2016. doi: <https://doi.org/10.3982/ECTA12144>. URL <https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA12144>.
- Alfonso Alba-Ramirez. Self-employment in the midst of unemployment: The case of spain and the united states. *Applied Economics*, 26(3):189–204, 1994.
- Elliott Ash and Stephen Hansen. Text algorithms in economics. *Annual Review of Economics*, 15(Volume 15, 2023):659–688, 2023. ISSN 1941-1391. doi: <https://doi.org/10.1146/annurev-economics-082222-074352>. URL <https://www.annualreviews.org/content/journals/10.1146/annurev-economics-082222-074352>.

- Patrick Bajari, Zhihao Cen, Victor Chernozhukov, Manoj Manukonda, Suhas Vijaykumar, Jin Wang, Ramon Huerta, Junbo Li, Ling Leng, George Monokroussos, and Shan Wan. Hedonic prices and quality adjusted price indices powered by ai, 2023.
- Beau Grant Barnes, Nancy L. Harp, and Derek Oler. Evaluating the sdc mergers and acquisitions database. *Financial Review*, 49(4):793–822, 2014a. doi: <https://doi.org/10.1111/fire.12057>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/fire.12057>.
- Beau Grant Barnes, Nancy L. Harp, and Derek Oler. Evaluating the sdc mergers and acquisitions database. *Financial Review*, 49(4):793–822, 2014b. doi: <https://doi.org/10.1111/fire.12057>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/fire.12057>.
- Kimberly Bayard, Emin Dinlersoz, Timothy Dunne, John Haltiwanger, Javier Miranda, and John Stevens. Early-stage business formation: An analysis of applications for employer identification numbers. Working Paper 24364, NBER, 2018.
- Anmol Bhandari and Ellen R. McGrattan. Sweat equity in u.s. private business. *Quarterly Journal of Economics*, 136(2):727–781, 2021.
- Anmol Bhandari, Paolo Martellini, and Ellen McGrattan. A theory of business transfers. *Working Paper*, March 2023.
- Christopher R. Bollinger, Barry T. Hirsch, Charles M. Hokayem, and James P. Ziliak. Trouble in the tails? what we know about earnings nonresponse 30 years after lillard, smith, and welch. *Journal of Political Economy*, 127(5):2143–2185, 2019.
- Francisco J. Buera. A dynamic model of entrepreneurship with borrowing constraints: Theory and evidence. *Annals of Finance*, 5(3-4):443–464, 2009.
- Marco Cagetti and Mariacristina DeNardi. Entrepreneurship, frictions, and wealth. *Journal of Political Economy*, 114(5):835–870, 2006.

- Sylvain Catherine. Keeping options open: What motivates entrepreneurs? *Journal of Financial Economics*, 144(1):1–21, 2022.
- Laurent Cavenaile, Murat Alp Celik, and Xu Tian. Are Markups Too High? Competition, Strategic Innovation, and Industry Dynamics. 2021.
- Laurent Cavenaile, Murat Alp Celik, and Xu Tian. The dynamic effects of antitrust policy on growth and welfare. *Journal of Monetary Economics*, 121:42–59, 2021. ISSN 0304-3932. doi: <https://doi.org/10.1016/j.jmoneco.2021.04.011>. URL <https://www.sciencedirect.com/science/article/pii/S0304393221000490>.
- Raj Chetty, Friedman John N., Emmanuel Saez, and Danny Yagan. The soi data-bank: A case study in leveraging administrative data in support of evidence-based policymaking. *Statistical Journal of the IAOS*, 34:99–103, 2018.
- Robert W. Crandall and Clifford Winston. Does antitrust policy improve consumer welfare? assessing the evidence. *Journal of Economic Perspectives*, 17(4):3–26, December 2003. doi: 10.1257/089533003772034871. URL <https://www.aeaweb.org/articles?id=10.1257/089533003772034871>.
- Colleen Cunningham, Florian Ederer, and Song Ma. Killer acquisitions. *Journal of Political Economy*, 129(3):649–702, 2021. doi: 10.1086/712506. URL <https://doi.org/10.1086/712506>.
- Joel M David. The Aggregate Implications of Mergers and Acquisitions. *The Review of Economic Studies*, 88(4):1796–1830, 11 2020. ISSN 0034-6527. doi: 10.1093/restud/rdaa077. URL <https://doi.org/10.1093/restud/rdaa077>.
- Angus Deaton. *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. World Bank, 1997.
- Jason DeBacker, Vasia Panousi, and Shanthi Ramnath. A risky venture: Income dynamics among pass-through business owners. *American Economic Journal: Macroeconomics*, 15(1):444–474, 2022.

- Ryan Decker, John Haltiwanger, Ron Jarmin, and Javier Miranda. The role of entrepreneurship in us job creation and economic dynamism. *Journal of Economic Perspectives*, 28(3):3–24, 2014.
- Larry G. Epstein and Stanley E. Zin. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica*, 57(4):937–969, 1989.
- David S. Evans and Boyan Jovanovic. An estimated model of entrepreneurial choice under liquidity constraints. *Journal of Political Economy*, 97(4):808–827, 1989.
- David S. Evans and Linda S. Leighton. Some empirical aspects of entrepreneurship. *American Economic Review*, 79(3):519–535, 1989.
- Robert W. Fairlie. Self-employment, entrepreneurship, and the nlsy79. *Monthly Labor Review*, pages 40–47, February 2005.
- Christian Fons-Rosen, Pau Roldan-Blanco, and Tom Schmitz. The aggregate effects of acquisitions on innovation and economic growth. 2021.
- Laurent Frésard, Gerard Hoberg, and Gordon M Phillips. Innovation Activities and Integration through Vertical Acquisitions. *The Review of Financial Studies*, 33(7):2937–2976, 09 2019. ISSN 0893-9454. doi: 10.1093/rfs/hhz106. URL <https://doi.org/10.1093/rfs/hhz106>.
- Alexandre Gaillard and Sumudu Kankanamge. Buying and selling entrepreneurial assets. *TSE Working Paper*, June 2020.
- Andrew Garin, Emilie Jackson, and Dmitri Koustas. New gig work or changes in reporting? understanding self-employment trends in tax data. Working Paper 2022–67, Becker Friedman Institute, 2022.
- Hoberg Gerard and Gordon M Phillips. Scope, scale and concentration: The 21st century firm. *Working Paper*, 2021.
- Robert E. Hall. Technical change and capital from the point of view of the dual. *Review of Economic Studies*, 35(1):35–46, 1968.

- Robert E Hall and Susan E Woodward. The burden of the nondiversifiable risk of entrepreneurship. *American Economic Review*, 100(3):1163–1194, 2010.
- Barton H. Hamilton. Does entrepreneurship pay? an empirical analysis of the returns to self-employment. *Journal of Political Economy*, 108(3):604–631, 2000.
- Amy Handlan. Text shocks and monetary surprises: Text analysis of fomc statements with machine learning. *Working Paper*, 2022.
- Anne Lundgaard Hansen and Sophia Kazinnik. Can chatgpt decipher fedpeak? *Working Paper*, 2023.
- Gerard Hoberg and Gordon Phillips. Product Market Synergies and Competition in Mergers and Acquisitions: A Text-Based Analysis. *The Review of Financial Studies*, 23(10):3773–3811, 08 2010. ISSN 0893-9454. doi: 10.1093/rfs/hhq053. URL <https://doi.org/10.1093/rfs/hhq053>.
- Erik Hurst and Annamaria Lusardi. Liquidity constraints, household wealth, and entrepreneurship. *Journal of Political Economy*, 112(2):319–347, 2004.
- Erik Hurst and Benjamin Wild Pugsley. What do small businesses do? *Brookings Papers on Economic Activity*, 2011(2):73–118, 2011.
- Boyan Jovanovic. Selection and the evolution of industry. *Econometrica*, 50(3): 649–670, 1982.
- Boyan Jovanovic and Peter L. Rousseau. The q-theory of mergers. *American Economic Review*, 92(2):198–204, May 2002. doi: 10.1257/000282802320189249. URL <https://www.aeaweb.org/articles?id=10.1257/000282802320189249>.
- Katya Kartashova. Private equity premium puzzle revisited. *American Economic Review*, 104(10):3297–3334, 2014.
- Sophia Kazinnik. Bank run, interrupted: Modeling deposit withdrawals with generative ai. *Working Paper*, 2023.

- Edward P. Lazear and Robert L. Moore. Incentives, productivity, and labor contracts. *Quarterly Journal of Economics*, 99(2):275–296, 1984.
- Ross Levine and Yona Rubinstein. Smart and illicit: Who becomes an entrepreneur and do they earn more? *Quarterly Journal of Economics*, 132(2):963–1018, 2016.
- Katherine Lim, Alicia Miller, Max Risch, and Eleanor Wilking. Independent contractors in the u.s.: New trends from 15 years of administrative tax data. SoI working paper, Internal Revenue Service, 2019.
- Jeremy Lise and Fabien Postel-Vinay. Multidimensional skills, sorting, and human capital accumulation. *American Economic Review*, 110(8):2328–2376, 2020.
- Gustavo Manso. Experimentation and the returns to entrepreneurship. *Review of Financial Studies*, 29(9):2319–2340, 2016.
- Ben Mermelstein, Volker Nocke, Mark A. Satterthwaite, and Michael D. Whinston. Internal versus external growth in industries with scale economies: A computational model of optimal merger policy. *Journal of Political Economy*, 128(1): 301–341, 2020. doi: 10.1086/704069. URL <https://doi.org/10.1086/704069>.
- Tobias J. Moskowitz and Annette Vissing-Jorgensen. The returns to entrepreneurial investment: A private equity premium puzzle? *American Economic Review*, 92(4):745–778, 2002.
- Simon C. Parker. *The Economics of Entrepreneurship*. Cambridge University Press, 2 edition, 2018.
- Jeremy G. Pearce and Liangjie Wu. Product Reallocation and Market Concentration. 2022.
- Michael Peters. Heterogeneous markups, growth, and endogenous misallocation. *Econometrica*, 88(5):2037–2073, 2020. doi: <https://doi.org/10.3982/ECTA15565>. URL <https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA15565>.

- Gordon M. Phillips and Alexei Zhdanov. R&D and the Incentives from Merger and Acquisition Activity. *The Review of Financial Studies*, 26(1):34–78, 01 2013. ISSN 0893-9454. doi: 10.1093/rfs/hhs109. URL <https://doi.org/10.1093/rfs/hhs109>.
- Vincenzo Quadrini. The importance of entrepreneurship for wealth concentration and mobility. *Review of Income and Wealth*, 45(1):1–19, 1999.
- Matthew Rhodes-Kropf and David T. Robinson. The market for mergers and the boundaries of the firm. *The Journal of Finance*, 63(3):1169–1211, 2008. doi: <https://doi.org/10.1111/j.1540-6261.2008.01355.x>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-6261.2008.01355.x>.
- Ellen Rissman. Self-employment as an alternative to unemployment. Working Paper 2003–34, Federal Reserve Bank of Chicago, 2003.
- Ellen Rissman. Labor market transitions and self-employment. Working Paper 2007–14, Federal Reserve Bank of Chicago, 2007.
- Matthew Smith, Danny Yagan, Owen Zidar, and Eric Zwick. Capitalists in the twenty-first century. *Quarterly Journal of Economics*, 134(4):1675–1745, 2019.
- George Tauchen. Finite state markov-chain approximations to univariate and vector autoregressions. *Economic Letters*, 20(2):177–181, 1986.
- U.S. Department of Justice and the Federal Trade Commission. Merger guidelines, 2023.

Appendix A

Innovation and Competition Policy

Derivations

Household Problem: The household's problem is

$$\begin{aligned} \max_{C_t} \mathbb{E} \int_0^\infty e^{-\rho t} \log C_t \, dt \\ s.t. \, \dot{A}_t = w_t + r_t A_t - C_t \\ C_t, H_t \geq 0. \end{aligned}$$

Demand: The final goods aggregator solves

$$\begin{aligned} \max_{Y_t, y_{i,t}} \quad & Y_t - \int_0^1 p_{i,t} y_{i,t} \, di \\ \text{s.t.} \quad & \log Y_t = \int_0^1 \log y_{i,t} \, di \\ & Y_t, y_{i,t} \geq 0. \end{aligned}$$

The first-order conditions result

$$p_{i,t} y_{i,t} = Y_t.$$

The market-level aggregator solves

$$\begin{aligned} \max_{y_{i,t}, y_{f,i,t}} \quad & p_{i,t} y_{i,t} - \sum_{f \in F_{i,t}} p_{f,i,t} y_{f,i,t} - p_{c,i,t} y_{c,i,t} \\ \text{s.t.} \quad & y_{i,t}^{\frac{\theta-1}{\theta}} = \sum_{f \in F_{i,t}} y_{f,i,t}^{\frac{\theta-1}{\theta}} + y_{c,i,t}^{\frac{\theta-1}{\theta}} \\ & y_{i,t}, y_{f,i,t} \geq 0. \end{aligned}$$

The first-order conditions result in market level demand of

$$y_{f,i,t} = p_{f,i,t}^{-\theta} p_{i,t}^{\theta} y_{i,t}.$$

The fact that the aggregator is competitive means

$$p_{i,t}^{1-\theta} = \sum_{f \in F_{i,t}} p_{f,i,t}^{1-\theta} + p_{c,i,t}^{1-\theta}.$$

Competition and Production: Within a market i , the firms $f \in F_{i,t}$ compete in Cournot competition. Taking their competitors' prices as given, the firm f solves

$$\begin{aligned} \max_{p_{f,i,t}, y_{f,i,t}, h_{f,i,t}, p_{i,t}} \quad & p_{f,i,t} y_{f,i,t} - w_t h_{f,i,t} \\ \text{s.t.} \quad & y_{f,i,t} = z_{f,t} q_{f,i,t} h_{f,i,t} \\ & y_{f,i,t} = p_{f,i,t}^{-\theta} p_{i,t}^{\theta-1} Y_t \\ & p_{i,t}^{1-\theta} = \sum_{f \in F_{i,t}} p_{f,i,t}^{1-\theta} + p_{c,i,t}^{1-\theta} \\ & y_{f,i,t}, h_{f,i,t} \geq 0. \end{aligned}$$

The problem is the same as

$$\begin{aligned} \max_{p_{f,i,t}, p_{i,t}} \quad & \left(p_{f,i,t} - \frac{w_t}{z_{f,t} q_{f,i,t}} \right) p_{f,i,t}^{-\theta} p_{i,t}^{\theta-1} Y_t \\ \text{s.t.} \quad & p_{i,t}^{1-\theta} = \sum_{f \in F_{i,t}} p_{f,i,t}^{1-\theta} + p_{c,i,t}^{1-\theta}. \end{aligned}$$

The first-order result in prices being pinned down by the system of equations

$$s_{f,i,t} = \frac{p_{f,i,t}^{1-\theta}}{\sum_{f' \in F_{i,t}} p_{f',i,t}^{1-\theta} + p_{c,i,t}^{1-\theta}}$$

$$\varepsilon_{f,i,t} = s_{f,i,t} + \theta(1 - s_{f,i,t})$$

$$p_{f,i,t} = \frac{\varepsilon_{f,i,t}}{\varepsilon_{f,i,t} - 1} \frac{w_t}{z_{f,t} q_{f,i,t}}.$$

Define the markup $m_{f,i,t}$ by

$$m_{f,i,t} = \frac{\varepsilon_{f,i,t}}{\varepsilon_{f,i,t} - 1}.$$

The $p_{i,t}$ becomes

$$p_{i,t}^{1-\theta} = \sum_{f \in F_{i,t}} \left(m_{f,i,t} \frac{w_t}{z_{f,t} q_{f,i,t}} \right)^{1-\theta} + \left(\frac{w_t}{\eta q_{i,t}} \right)^{1-\theta}$$

where $q_{i,t} = \max_{f \in F_{i,t}} q_{f,i,t}$. Define $\tilde{q}_{f,i,t} = \frac{q_{f,i,t}}{q_{i,t}}$. This becomes

$$p_{i,t} = \left(\sum_{f \in F_{i,t}} \left(\frac{m_{f,i,t}}{z_{f,t} \tilde{q}_{f,i,t}} \right)^{1-\theta} + \left(\frac{w_t}{\eta} \right)^{1-\theta} \right)^{\frac{1}{1-\theta}} \frac{w_t}{q_{i,t}}$$

Define

$$\begin{aligned}
z_{i,t} &= \max\{z_{f,t} : f \in F_{i,t}\}, \\
\bar{z}_{f,i,t} &= \frac{z_{f,t}}{z_{i,t}}, \\
m_{i,t} &= \left(\sum_{f \in F_{i,t}} \left(\frac{m_{f,i,t}}{\bar{z}_{f,i,t} \tilde{q}_{f,i,t}} \right)^{1-\theta} + \left(\frac{1}{\eta z_{i,t}} \right)^{1-\theta} \right)^{\frac{1}{1-\theta}}.
\end{aligned}$$

Market-level prices are then

$$p_{i,t} = m_{i,t} \frac{w_t}{q_{i,t} z_{i,t}}.$$

Output, labor demand, and profits are

$$\begin{aligned}
y_{f,i,t} &= m_{f,i,t}^{-\theta} \tilde{q}_{f,i,t}^{\theta} \tilde{z}_{f,i,t}^{\theta} m_{i,t}^{\theta-1} q_{i,t} z_{i,t} \frac{Y_t}{w_t} \\
h_{f,i,t} &= m_{f,i,t}^{-\theta} \tilde{q}_{f,i,t}^{\theta-1} \tilde{z}_{f,i,t}^{\theta-1} m_{i,t}^{\theta-1} \frac{Y_t}{w_t} \\
\pi_{f,i,t} &= (m_{f,i,t}^{1-\theta} - m_{f,i,t}^{-\theta}) \tilde{q}_{f,i,t}^{\theta-1} \tilde{z}_{f,i,t}^{\theta-1} m_{i,t}^{\theta-1} Y_t.
\end{aligned}$$

For the competitive firms, output and labor demand is

$$\begin{aligned}
y_{c,i,t} &= \eta^{\theta} q_{i,t} z_{i,t}^{1-\theta} m_{i,t}^{\theta-1} \frac{Y_t}{w_t} h_{c,i,t} &= \eta^{\theta-1} z_{i,t}^{1-\theta} m_{i,t}^{\theta-1} \frac{Y_t}{w_t}
\end{aligned}$$

Goods Aggregation: Market-level output $y_{i,t}$ is

$$\begin{aligned} y_{i,t} &= \left(\sum_{f \in F_{i,t}} \left(m_{f,i,t}^{-\theta} m_{i,t}^{\theta-1} z_{f,t}^{\theta} \tilde{q}_{f,i,t}^{\theta} q_{i,t} z_{i,t} \frac{Y_t}{w_t} \right)^{\frac{\theta-1}{\theta}} + \left(\eta^{\theta} z_{i,t}^{\theta} m_{i,t}^{\theta-1} q_{i,t} \frac{Y_t}{w_t} \right)^{\frac{\theta-1}{\theta}} \right) \\ &= \frac{q_{i,t} z_{i,t}}{m_{i,t}} \frac{Y_t}{w_t}. \end{aligned}$$

Applying this to the output aggregator yields

$$\log Y_t = \int_0^1 \log \left(\frac{q_{i,t} z_{i,t}}{m_{i,t}} \frac{Y_t}{w_t} \right) di$$

and

$$w_t = Q_t Z_t \exp \left(\int_0^1 \log m_{i,t}^{-1} di \right)$$

where

$$\begin{aligned} Q_t &= \exp \left(\int_0^1 \log q_{i,t} di \right) \\ Z_t &= \exp \left(\int_0^1 \log z_{i,t} di \right). \end{aligned}$$

Labor Aggregation: Market-level labor demand is

$$\begin{aligned} h_{i,t} &= \sum_{f \in F_{i,t}} h_{f,i,t} + h_{c,i,t} \\ &= \left(\sum_{f \in F_{i,t}} m_{f,i,t}^{-\theta} z_{f,t}^{\theta-1} \tilde{q}_{f,i,t}^{\theta-1} + \eta^{\theta-1} \right) m_{i,t}^{\theta-1} \frac{Y_t}{w_t} \end{aligned}$$

It follows that aggregate labor demand is

$$H_t = \int_0^1 \left(\sum_{f \in F_{i,t}} m_{f,i,t}^{-\theta} z_{f,t}^{\theta-1} \tilde{q}_{f,i,t}^{\theta-1} + \eta^{\theta-1} \right) m_{i,t}^{\theta-1} \frac{Y_t}{w_t} di.$$

Output: Plugging the wage condition into aggregate labor demand gives

$$Y_t = Q_t Z_t \mathcal{M}_t$$

where

$$\mathcal{M}_t = \frac{\exp \left(\int_0^1 \log m_{i,t}^{-1} di \right)}{\int_0^1 \left(\sum_{f \in F_{i,t}} m_{f,i,t}^{-\theta} \tilde{z}_{f,t}^{\theta-1} \tilde{q}_{f,i,t}^{\theta-1} + \left(\frac{\eta}{z_{i,t}} \right)^{\theta-1} \right) m_{i,t}^{\theta-1} di}$$

Growth Rate: On the balanced-growth path, the growth rate is

$$\begin{aligned}
g &= \frac{d}{dt} \log w_t \\
&= \underbrace{\frac{d}{dt} \log Q_t}_{\text{Change in Goods Productivity}} + \underbrace{\frac{d}{dt} \log Z_t}_{\text{Change in Firm}} + \underbrace{\frac{d}{dt} \int_0^1 \log m_{i,t}^{-1} di}_{\text{Change in Markups}}.
\end{aligned}$$

Beginning with the change in goods' productivity portion, let $\Delta > 0$.

$$\begin{aligned}
\frac{\log Q_{t+\Delta} - \log Q_t}{\Delta} &= \int_0^1 \sum_{f \in F_{i,t}: \tilde{q}_{f,i,t}=1} \log(\lambda) j_{f,i,t} di + \log(\lambda) \sum_z X_{z,t} \\
&\quad + \int_0^1 \sum_{\substack{f', f \in F_{i,t} \\ f \neq f'}} \mathbb{E} \left[\log \left(\frac{q_{\ell, f', f, i, t}^{MA}}{q_{\ell, i, t}} \right) \right] a_{f', f, i, t} (1 - B_{f', f, i, t}) di + O(\Delta)
\end{aligned}$$

$q_{\ell, f', f, i, t}^{MA}$ is the market leader's good's productivity after firm f acquires the good of firm f' . That means

$$\begin{aligned}
\frac{d}{dt} \log Q_t &= \underbrace{\int_0^1 \sum_{f \in F_{i,t}: \tilde{q}_{f,i,t}=1} \log(\lambda) j_{f,i,t} di}_{\text{Growth Through Innovation}} + \underbrace{\log(\lambda) \sum_z X_{z,t}}_{\text{Growth Through Entry}} \\
&\quad + \underbrace{\int_0^1 \sum_{\substack{f', f \in F_{i,t} \\ f \neq f'}} \mathbb{E} \left[\log \left(\frac{q_{\ell, f', f, i, t}^{MA}}{q_{\ell, i, t}} \right) \right] a_{f', f, i, t} (1 - B_{f', f, i, t}) di}_{\text{Growth Through M\&A}}.
\end{aligned}$$

Moving on to the the other two terms, they depend only on the firms' productivity and goods' relative productivity in market i . Only the balanced growth path, the

distribution of these productivities is constant. Thus, there is no change in markup over time.

Appendix B

Numerical Solution Method for the Innovation and Competition Policy Model

B.1 High-Level Overview

While the model appears quite complicated, the model is just a combination of many simple parts. This appendix describes how to solve the model numerically. Firm policies are computed by combining a matrix-vector product and element-wise operations. Updating the value function is solving a linear system. This linear system is strictly diagonally dominant, meaning it can be solved by an iterative

method (in this case, Jacobi's method). This makes the value function problem tractable. Computing the market distribution is another series of matrix-vector products. Finally, prices are updated by another set of matrix-vector products and element-wise operations.

B.2 Model Solution Steps

B.2.1 Construct the State Set

Market Set: Each division in a market is characterized by firm-level productivity z and steps behind good-level productivity level Δ . The market state is the list of these values for every firm operating in the market. A monopoly market will have only one tuple $\{(z_1, \Delta_1)\}$. A market can have at most \bar{F} firms $\{(z_1, \Delta_1), \dots, (z_{\bar{F}}, \Delta_{\bar{F}})\}$. The market set is every combination of $z \in \{z_1, z_2, z_3\}$, $\Delta \in 0, 1, \dots, \bar{n}$, and number of firms in the market from 1 to \bar{F} . There are multiple symmetries in firm states – the same set of firms should only be included once. Enumerate every market state to create a unique market index.

Division Set: A division's state is its own (z, Δ) as well as the market in which the firm operates. Since multiple firms in a market can have the same (z, Δ) , that type of firm must only be included once. Enumerate every division state to create a

unique division index.

Transaction Set: It is also helpful to have a set of every possible M&A transaction in the economy. For every market, a transaction is a combination of a buyer division, a seller division, and a set of uninvolved divisions. Enumerate every transaction to create a unique transaction index.

B.2.2 Solve the Cournot Game

Given the parameter set and the market state set, the goal is to solve the Cournot game for every market state. The key output is division-level profits and market share.

Input: The parameter set \mathcal{P} and market state set \mathcal{S}_M .

1. For each market state $((z_1, \Delta_1), \dots, (z_N, \Delta_N))$ in \mathcal{S}_M .
 - (a) Compute the good's specific productivity $q_k = \lambda^{-\Delta_k}$ for $k \in 1, 2, \dots, N$
 - (b) Compute the fringe price by

$$p_c = \frac{w}{\eta \max\{q_k\}}$$

- (c) Solve for prices p_k , market shares s_k , and elasticities ε_k for $k \in 1, 2, \dots, N$

by solving the nonlinear system

$$s_k = \frac{p_k^{1-\theta}}{\sum_{k'=1}^N p_{k'}^{1-\theta} + p_c^{1-\theta}}$$

$$\varepsilon_k = \theta + (1 - \theta)s_k$$

$$p_k = \frac{\varepsilon_k}{\varepsilon_k - 1} \frac{w}{z_k q_k}$$

(d) Compute the market-level price index p by

$$p = \left(\sum_{k=1}^N p_k^{1-\theta} + p_c^{1-\theta} \right)^{\frac{1}{1-\theta}}.$$

(e) Compute division-level output y_k by

$$y_k = p_k^{-\theta} p^{\theta-1}.$$

(f) Compute division-level labor supply h_k by

$$h_k = \frac{y_k}{z_k q_k}.$$

(g) Compute division-level profits π_k by

$$\pi_k = p_k y_k - w h_k.$$

(h) Compute market-level hhi by

$$hhi = \sum_{k=1}^N s_k^2.$$

B.2.3 Construct Division Linear Maps

The goal is to convert the division value function problem (1.27) into a series of simple optimality conditions and a linear system. The linear system is $A^D \times D = b^D$ where A^D defined by

$$\begin{aligned} A^D = & \rho I \\ & - \sum_{n=1}^{\bar{F}} \text{diag}(M_n^{J,D} \times j) \times A_n^{J,D} \\ & - \sum_{z \in Z} X_z \times A_z^{X,D} \\ & - \sum_{(n,m) \in MA} \text{diag}(M_{n,m}^{\text{MA},D} \times a^{\text{combine}}) \times A_{n,m}^{\text{combine},D} \\ & - \sum_{(n,m) \in MA} \text{diag}(M_{n,m}^{\text{MA},D} \times a^{\text{kill}}) \times A_{n,m}^{\text{kill},D}. \end{aligned}$$

This section will describe each of the matrices that build A^D .

Optimal Improvement R&D Effort A_{effort}^J : For a division, their optimal improvement condition is given by

$$\varphi_J(1 + \gamma_J)j^{\gamma_J} = D_{f,i,t}(q_{f,i,t} \rightarrow \lambda q_{f,i,t}) - D_{f,i,t}.$$

The right-hand side is the value of the division having an improved product minus the current value of the division. Since the solution method is targeted for GPU computation, the goal is to encode every division state's right-hand side in a sparse matrix denoted A_{effort}^J . Create a sparse matrix where the number of rows and columns is the number of division states. Let i be the division's index before improvement and i' division's index after improving. Set the (i, i) index of A_{effort}^J to -1 and the (i, i') index of A_{effort}^J to 1 . The optimality condition is

$$\varphi_J(1 + \gamma_J)j^{\gamma_J} = A_{\text{effort}}^J D$$

where D is now the vector of every divisions' value. The vector of optimal for every division j can then be computed by a sparse matrix-vector product and two element-wise operations.

Improvement R&D Results $A_n^{J,D}$: The division's value function tracks the impact

of both the division and its competitors improving by

$$\begin{aligned}
& \underbrace{j_{f,i,t} (D_{f,i,t} (q_{f,i,t} \rightarrow \lambda q_{f,i,t}) - D_{f,i,t})}_{\text{Firm's R\&D Improvement}} \\
& + \underbrace{\sum_{\substack{f' \in F_{i,t} \\ f' \neq f}} j_{f',i,t} (D_{f,i,t} (q_{f',i,t} \rightarrow \lambda q_{f',i,t}) - D_{f,i,t})}_{\text{Competitors' R\&D Improvement}}
\end{aligned}$$

The differences can be represented as a set of \bar{F}^1 sparse matrices denoted $A_n^{J,D}$ where n denotes the n th firm in the market. Let $A_1^{J,D}$ be the matrix capturing the division's own improvement. This matrix is the same as A_{effort}^J . What remains to be constructed is the matrices for improvement by the competitors. These are constructed in the same way as A_{effort}^J ; the only difference is the new index is now the index if one of the firm's competitors improves their goods. If a market has less than n firms in its market, then the row of the sparse matrix is left zero.

Improvement Effort Map $M_n^{J,D}$: Each division state has one entry in the j vector. However, each division's j needs to be accounted for in both its own value function as well as all of its competitors. We need to generate a map from j to the appropriate row of $A_n^{J,D}$. For row i of $A_n^{J,D}$, let k be the division index of the improving division. In $M_n^{J,D}$, set (i, k) equal to one.

¹the maximum number of firms allowed in a market.

Entry Results $A_z^{X,D}$: The division's value function also tracks entry with the term

$$\underbrace{\sum_z x_{z,t} (D_{f,i,t}(\{q_{f',i,t}\} \rightarrow \{q_{f',i,t}\} \cup q_{f'',i,t}) - D_{f,i,t})}_{\text{New Competitor's Entry}}.$$

The entry result matrix $A_z^{X,D}$ tracks the difference in division values from this equation. Let i denote the division's index and i' denote the division's index after a firm with productivity z enters into the market. There are two cases. If division i remains in the market, then set the (i, i) entry to -1 and the (i, i') entry to 1 . If the entry of the new firm results in division i exiting the market, then just set (i, i) to -1 .

Combining Surplus Matrix A_{TS}^{combine} : This matrix tracks the total surplus from a M&A transaction where the products are combined. If a transaction's combined productivity lands in the middle of two rungs on the quality ladder, it is randomly assigned to one of those two rungs. Let q be the resulting combined productivity, q_h be the rung on the quality ladder immediately above q , and q_l be the rung on the quality ladder immediately below q . Define

$$\omega = \frac{q - q_l}{q_h - q_l}.$$

The probability of the transaction being assigned to q_h is ω . Let k denote the transaction's index, k_b be the division index of the buyer, k_s be the division index of the

seller, k_h be the high index of the combined product, and k_l be the low index of the combined product. The total surplus of the transaction

$$\omega D_{k_h, i, t} + (1 - \omega) D_{k_l, i, t} - D_{k_b, i, t} - D_{k_s, i, t}$$

To defined A_{TS}^{combine} , set the entries (k, k_b) and (k, k_s) to -1 , (k, k_h) to ω , and (k, k_l) to $1 - \omega$.

Combining Result Matrix $A_{n, m}^{\text{combine}, D}$: The division value function tracks the impact of M&A transaction between competitor divisions

$$\underbrace{\sum_{\substack{f', f'' \in F_{i, t} \\ f', f'' \neq f \\ f' \neq f''}} a_{f'', f', i, t} ((1 - B_{f'', f', i, t}) \mathbb{E}[D_{f, i, t}(MA_{f'', f', i, t}(\varepsilon_{f'', f', i, t}))] - D_{f, i, t})}_{\text{Competitors' M\&A Actions}}.$$

The $A_{n, m}^{\text{kill}, D}$ tracks the impact of transactions that result in combination. Let k denote the division's index prior to the transaction, k_h denote the high index, k_l denote the low index, and ω be the probability of the resulting transaction moving to the high index. Set (k, k) to -1 , (k, k_h) to ω , and (k, k_l) to $1 - \omega$.

Killing Surplus Matrix A_{TS}^{kill} : This matrix tracks the total surplus from a M&A transaction where the buyer kills the seller's product. Let k denote the transaction's index, k_b denote the buyer's division's index, k_s denote the seller's division index, and k'_b denote the buyer's division's index after the transaction is completed. The

total surplus of this transaction is

$$D_{k'_b, i, t} - D_{k_b, i, t} - D_{k_s, i, t}.$$

To compute M&A effort, we need to compute this surplus. Define the surplus map

A_{TS}^{kill} by setting (k, i_b) and (k, i_s) to -1 and (k, i'_b) to 1 for every k .

Killing Result Matrix $A_{n,m}^{\text{kill}, D}$: The division value function tracks the impact of M&A transaction between competitor divisions

$$\underbrace{\sum_{\substack{f', f'' \in F_{i,t} \\ f', f'' \neq f \\ f' \neq f''}} a_{f'', f', i, t} ((1 - B_{f'', f', i, t}) \mathbb{E}[D_{f, i, t}(MA_{f'', f', i, t}(\varepsilon_{f'', f', i, t}))] - D_{f, i, t})}_{\text{Competitors' M\&A Actions}}.$$

The $A_{n,m}^{\text{kill}, D}$ tracks the impact of transactions that result in killer acquisitions. Let k denote the division's index prior to the transaction and k' denote the index after the transaction. Set (k, k) to -1 and (k, k') to 1 .

M&A Map $M_{n,m}^{MA, D}$: This is a similar map to $M_n^{J, D}$. It maps from a transaction index to the appropriate row of the $A_{n,m}^{\text{combine}, D}$ and $A_{n,m}^{\text{kill}, D}$ matrices. Build $M_{n,m}^{MA, D}$ by setting (i, k) to 1 where i is the division index and k is there (n, m) 's transaction index.

B.2.4 Construct Market Linear Maps

The next goal is to be able to write the Kolmogorov forward equation (1.28) as

$$\dot{\mu}_t = A^M \mu$$

where

$$\begin{aligned} A^M &= \sum_{n=1}^{\bar{F}} A_n^{J,M} \times \text{diag} (M_n^{J,M} \times j) \\ &+ \sum_{z \in Z} X_z \times A_z^{X,M} \\ &+ \sum_{(n,m) \in MA} A_{n,m}^{\text{combine},M} \times \text{diag} (M_{n,m}^M \times a^{\text{combine}}) \\ &+ \sum_{(n,m) \in MA} A_{n,m}^{\text{kill},M} \times \text{diag} (M_{n,m}^M \times a^{\text{kill}}). \end{aligned}$$

This subsection will describe each of the matrices that build A^M .

R&D Improvement Result $A_n^{J,M}$: Represents how a market will change when firm n has a successful innovation. In the Kolmogorov equation, this term is

$$\begin{aligned} &- \sum_{k=1}^{n(S_t)} j_k(S_t) \mu_t(S_t) \\ &+ \int_{\{S'_t: \text{Firm } k \text{ innovation means } S'_t \text{ becomes } S_t\}} j_k(S'_t) d\mu_t(S'_t). \end{aligned}$$

Let i denote the market index before the n th firm innovates and i' denote the market index after the innovation. Within $A_n^{J,M}$, set the index (i, i) to -1 and (i', i) to 1 .

R&D Effort Map $M_n^{J,M}$: Maps from the innovating division index to the market index. Let i denote the market index and f be the division index of the n th firm in the market. Set the (i, f) entry in $M_n^{J,M}$ to one.

Entry Result $A_z^{X,M}$: Represents how a market will change when a new firm with productivity z enters into the market. In the Kolmogorov equation, this is

$$- \sum_z x_{z,t} \mu_t(S_t).$$

Let i denote the market index before entry of a firm with productivity z and i' denote the market index after entry. Set the index (i, i) to -1 and (i', i) to 1 .

M&A Combine Result $A_{i,j}^{\text{combine},M}$: Represents how a market will change when division i buys division j and combines with division j 's product. In the Kolmogorov equation, this is part of

$$\begin{aligned} & - \sum_{k=1}^{n(S_t)} \sum_{k' \neq k} a_{k',k}(S_t) \mu_t(S_t) \\ & + \int_{\{S'_t: \text{Firm } k' \text{ and } k \text{ merge means } S'_t \text{ becomes } S_t\}} (1 - B_{k',k,t}(S'_t)) a_{k',k,t}(S'_t) d\mu_t(S'_t). \end{aligned}$$

Let k denote the market index before the transaction, k_h denote the high-index after

combining, k_l denote the low-index after combining, and ω denote the probability of high index. Set the index (k, k) to -1 , (k_h, k) to ω , (k_l, k) to $1 - \omega$.

M&A Kill Result $A_{i,j}^{\text{kill},M}$: Represents how a market will change when division i buys division j and kills division j 's product. In the Kolmogorov equation, this is also part of

$$\begin{aligned}
& - \sum_{k=1}^{n(S_t)} \sum_{k' \neq k} a_{k',k}(S_t) \mu_t(S_t) \\
& + \int_{\{S'_t: \text{Firm } k' \text{ and } k \text{ merge means } S'_t \text{ becomes } S_t\}} (1 - B_{k',k,t}(S'_t)) a_{k',k,t}(S'_t) d\mu_t(S'_t).
\end{aligned}$$

Let k denote the market index before the transaction and k' denote the market index after the transaction. Set the index (k, k) to -1 and (k', k) to 1 .

M&A Map $M_{i,j}^M$: Maps from a transaction index to the market index. Let k denote the market index and let k' be the transaction index of (i, j) transaction in market k . Set (k, k') to one.

B.2.5 Construct Entry Map

The last map we need is a map to compute $\mathbb{E}[D|z]$ for entry decision-making. This operator will be approximated by

$$(M_z^X \times D) \cdot \mu.$$

This subsection describes the entry map M_z^X . This matrix is the number of productivity states by the number of division states. For every market i and entry productivity z , let i_d denote the division index of the new firm. Let i_z be the productivity index for z . Set (i_z, i_d) to one.

B.2.6 Computing Antitrust Blocking Probabilities

For each potential transaction, let hhi be the HHI before the transaction, hhi' be the HHI after the transaction, and $\Delta hhi = hhi' - hhi$. Then compute the blocking probability according to (1.49).

B.2.7 Solve for the Stationary Distribution

Starting with a candidate division value D , improvement R&D effort j , combine M&A effort ma^{effort} , kill M&A effort ma^{kill} , market distribution μ , and entry rate $\{X_z\}$

1. Solve for the Division Value Function

(a) Compute Improvement R&D Effort

- i. Compute the R&D division improvement effort by computing ΔD_J

$$\Delta D_J = \max\{0, A_{\text{effort}}^J \times D\}$$

and then computing improvement effort j

$$j = \left(\frac{\Delta D_J}{\varphi_J(1 + \gamma_J)} \right)^{\frac{1}{\gamma_J}}.$$

ii. Compute improvement R&D cost c_J by

$$c^J = \varphi_J j^{1+\gamma_J}.$$

(b) Compute M&A Effort

i. Compute the combining and killing M&A total surplus by

$$TS^{\text{combine}} = \max\{0, A_{TS}^{\text{combine}} \times D\}$$

$$TS^{\text{kill}} = \max\{0, A_{TS}^{\text{kill}} \times D\}$$

ii. Compute the gain from combining products

$$\Delta TS = \max\{0, TS^{\text{combine}} - TS^{\text{kill}}\}.$$

iii. Compute the probability of combining products

$$\mathbb{P}[\text{combine}] = 1 - e^{-\omega \Delta TS}.$$

iv. Compute the expected value of the combination shock

$$\mathbb{E}[\varepsilon] = \frac{1}{\omega} \left(1 - (\omega \mathbb{P}[\text{combine}] + 1) e^{-\omega \mathbb{P}[\text{combine}]} \right)$$

v. Compute the probability of the transaction being blocked

$$\mathbb{P}[\text{block}] = \mathbb{P}[\text{block}|\text{combine}] \mathbb{P}[\text{combine}] + \mathbb{P}[\text{block}|\text{kill}] (1 - \mathbb{P}[\text{combine}])$$

vi. Compute the expected total surplus of the transaction

$$\mathbb{E}[TS] = \max \left\{ 0, \mathbb{P}[\text{combine}] TS^{\text{combine}} - \mathbb{E}[\varepsilon] + (1 - \mathbb{P}[\text{combine}]) TS^{\text{kill}} \right\}.$$

vii. Compute the expected buyer's and seller's share of the surplus

$$\mathbb{E}[BS] = \beta (1 - \mathbb{P}[\text{block}]) \mathbb{E}[TS]$$

$$\mathbb{E}[SS] = (1 - \beta) (1 - \mathbb{P}[\text{block}]) \mathbb{E}[TS]$$

viii. Compute M&A effort a by

$$a = \left(\frac{\mathbb{E}[BS]}{\varphi_a(1 + \gamma_a)} \right)^{\frac{1}{\gamma_a}}.$$

ix. Compute combine and kill M&A arrival rates by

$$a^{combine} = (1 - \mathbb{P}[block]) \mathbb{P}[combine] a$$

$$a^{kill} = (1 - \mathbb{P}[block]) (1 - \mathbb{P}[combine]) a$$

x. Compute M&A cost

$$c^{MA} = \varphi_a a^{1+\gamma_a}.$$

xi. Compute M&A flow by

$$ma^{\text{flow}} = M_{buyer}^{ma,D} \times (\mathbb{E}[BS] a) + M_{seller}^{ma,D} \times (\mathbb{E}[SS] a).$$

(c) **Update D**

i. Compute the linear system's matrix A^D by

$$\begin{aligned}
A^D &= \rho I \\
&- \sum_{n=1}^{\bar{F}} \text{diag}(M_n^{J,D} \times j) \times A_n^{J,D} \\
&- \sum_{z \in Z} X_z \times A_z^{X,D} \\
&- \sum_{(n,m) \in MA} \text{diag}(M_{n,m}^{\text{MA},D} \times a^{\text{combine}}) \times A_{n,m}^{\text{combine},D} \\
&- \sum_{(n,m) \in MA} \text{diag}(M_{n,m}^{\text{MA},D} \times a^{\text{kill}}) \times A_{n,m}^{\text{kill},D}.
\end{aligned}$$

ii. Compute the right-hand side of the linear system b_D by

$$b^D = \pi - c^J - c^{MA} + ma^{\text{flow}}$$

iii. Solve the linear system (A^D, b^D) by Jacobi's method. This results in
a new division value function D^{new}

iv. Update D by

$$D = \text{step} \times D^{\text{new}} + (1 - \text{step}) \times D$$

for some step size in $[0,1]$

(d) Repeat steps (a)-(d) until convergence of D .

2. Compute The Market Distribution

- (a) Compute the Komologrov equation matrix A^M by

$$\begin{aligned}
 A^M &= \sum_{n=1}^{\bar{F}} A_n^{J,M} \times \text{diag} (M_n^{J,M} \times j) \\
 &+ \sum_{z \in Z} X_z \times A_z^{X,M} \\
 &+ \sum_{(n,m) \in MA} A_{n,m}^{\text{combine},M} \times \text{diag} (M_{n,m}^M \times a^{\text{combine}}) \\
 &+ \sum_{(n,m) \in MA} A_{n,m}^{\text{kill},M} \times \text{diag} (M_{n,m}^M \times a^{\text{kill}})
 \end{aligned}$$

- (b) Discretize time dimensions of the Komologrov equation

$$\frac{\mu_{t+h} - \mu_t}{h} = A^M \mu_t$$

for some $h > 0$.

- (c) Update the market distribution according to

$$\mu_{t+h} = \mu_t + h A^M \mu_t.$$

- (d) Repeat step (c) until the distribution converges.

3. Update Entry Rates

- (a) For each $z \in Z$, compute the expected value of entry

$$\mathbb{E} [D^{\text{entry}}|z] = (M_z^X \times D) \cdot \mu.$$

- (b) Compute updated entry rates

$$X_z^{\text{new}} = \left(\frac{\mathbb{E} [D^{\text{entry}}|z]}{\varphi_X(1 + \gamma_X)} \right)^{\frac{1}{\gamma_X}}.$$

- (c) Update the entry rate

$$X_z = \text{step} \times X_z^{\text{new}} + (1 - \text{step}) \times X_z.$$

4. Repeat step 1-3 until convergence of the entry rates.

B.3 Supporting Proofs

In the division value function solver, the linear system was solved by Jacobi's method.

This method requires that the linear system is strictly diagonally dominant.

Lemma 5. *The matrix A^D is strictly diagonally dominant.*

Proof. The diagonal terms of A^D are

$$\begin{aligned}
\mathbf{diagonal} &= \rho \\
&+ \sum_{n=1}^{\bar{F}} M_n^{J,D} \times j \\
&+ \sum_{z \in Z} X_z \\
&+ \sum_{(i,j) \in MA} M_{i,j}^{\text{combine},D} \times ma^{\text{combine}} \\
&+ \sum_{(i,j) \in MA} M_{i,j}^{\text{kill},D} \times ma^{\text{kill}}
\end{aligned}$$

The sum of the absolute value of off-diagonals is

$$\begin{aligned}
\mathbf{off-diagonal} &= \sum_{n=1}^{\bar{F}} \left| -M_n^{J,D} \times j \right| \\
&+ \sum_{z \in Z} \left| -X_z \right| \\
&+ \sum_{(i,j) \in MA} \left| -M_{i,j}^{\text{combine},D} \times ma^{\text{combine}} \right| \\
&+ \sum_{(i,j) \in MA} \left| -M_{i,j}^{\text{kill},D} \times ma^{\text{kill}} \right|
\end{aligned}$$

Given that $j \geq 0$, $X \geq 0$, $a^{\text{combine}} \geq 0$, and $a^{\text{kill}} \geq 0$, we know

$$\mathbf{diagonal} = |\mathbf{diagonal}|$$

and

$$\text{diagonal} - \text{off-diagonal} = \rho.$$

Since $\rho > 0$, A^D is strictly diagonally dominant. ■

Appendix C

Antitrust Policy Regression

This appendix provides a deeper description of the logistic regression used in Section 1.3 to estimate $\hat{\mathbb{P}}(B|S, \text{hhi}, \Delta\text{hhi})$.

C.1 Data Preparation

The data source for this regression is the Federal Trade Commission publication *Horizontal Mergers Investigation Data, Fiscal Years 1996-2011*. Table 3.1 of the publication details the outcome of 1435 transactions that underwent a second review. Approximately 76% of these transactions resulted in an enforcement action. The average post-merger HHI of these transactions is 0.4609 and the average change in HHI is 0.1573.

One limitation of the data is that the report does not directly report each transaction’s post-merger HHI and change in HHI. Instead, it bins transactions into ranges of post-merger HHI and changes in HHI¹. Logistic regression requires numeric values to estimate the blocking probability. To convert the bins to numeric values, the midpoint of the bin is selected for all transactions within the bin². Two sets of bins require special consideration. One set is transactions where the post-merger HHI is ”0.7+”. The upper range of this bin is taken to be 1.0 (the maximum possible HHI), meaning the midpoint is 0.85. The second is the set of transactions with a change in HHI greater than 0.2. The upper range for the change in HHI is set at 0.5.

C.2 Regression Results

The logistic regression model is

$$\text{logit } \hat{\mathbb{P}}(B|S, \text{hhi}, \Delta\text{hhi}) = \beta_0 + \beta_{\text{hhi}} \text{hhi} + \beta_{\Delta\text{hhi}} \Delta\text{hhi} + \beta_{\text{hhi}, \Delta\text{hhi}} \text{hhi} \times \Delta\text{hhi} + \varepsilon. \quad (\text{C.1})$$

The estimation results are shown in Table C.1 and the resulting conditional probability function is displayed in Figure C.1. The resulting blocking probability increases with both hhi and Δhhi . For transactions in the review window recommended for

¹See Table 1.1 for a replication of the data.

²An alternative approach is sampling uniformly within the bin. However, this is undesirable as it adds additional noise to the data that can lower the model’s predictive power

review in the current horizontal merger investigation guidelines (a post-merger HHI above 0.15 and change in HHI above 0.05), the average blocking probability conditional on second review is approximately 87%.

C.3 Tables

Table C.1: $\hat{\mathbb{P}}(B|S, \text{hhi}, \Delta\text{hhi})$ Regression Results

	logit $\hat{\mathbb{P}}(B S, \text{hhi}, \Delta\text{hhi})$
Constant	−0.216
hhi	1.862**
Δhhi	0.186
$\text{hhi} \times \Delta\text{hhi}$	10.342**

C.4 Figures

Figure C.1: Estimated $\hat{\mathbb{P}}(B|S, \text{hhi}, \Delta \text{hhi})$.

