### ABSTRACT

Title of dissertation: ESSAYS ON FIRM DYNAMICS AND

JOB CREATION

Joonkyu Choi

Doctor of Philosophy, 2018

Dissertation directed by: Professor S. Boragan Aruoba

Professor John Haltiwanger Department of Economics

I describe two studies on firm dynamics and job creation. In Chapter 1, I identify a key predictor of the early growth trajectory of young firms: the outside options of the business founders. I show that entrepreneurs with higher outside options as paid workers tend to take larger business risks, and thus exhibit a more up-or-out type of firm dynamics. I find empirical support for the model's predictions using a large founder-firm matched data set built from administrative databases of the U.S. Census Bureau. I find that controlling for past business performance, young firms operated by entrepreneurs with higher outside options exhibit (i) higher firm exit rates, (ii) more growth dispersion, and (iii) faster growth conditioning on survival. With the calibrated model, I find that deterioration in the outside options of entrepreneurs can have a sizable negative impact on aggregate output and productivity via lower risk-taking by young firms and slower growth in their life cycle. These findings indicate that the expected post-failure outcomes of entrepreneurs are an important factor that governs young firm growth as well as aggregate output and

productivity.

Chapter 2 studies how firms' lobby behavior affects the allocation of federal procurement contracts during the fiscal stimulus period and the magnitude of the local job creation effect. Using the allocation of contracts under the American Recovery and Reinvestment Act (ARRA) as a laboratory, it is shown that among firms with a similar propensity to lobby on ARRA in 2009, those firms that actually lobbied on ARRA-related bills won 5% more contracts and 50% larger contracts than firms that did not. We further investigate whether procurement spending channeled through lobbying firms has a differential impact on MSA-level employment growth. We find that \$1 million procurement spending yields on average of 11.5 jobs, and that the effect is entirely driven by contracts channeled through non-lobbying firms. While procurement channeled through lobbying firms has no significant impact on job creation, \$1 million in procurement spending through non-lobbying firms yields 16 jobs.

# ESSAYS ON FIRM DYNAMICS AND JOB CREATION

by

### Joonkyu Choi

Dissertation submitted to the Faculty of the Graduate School of the University of Maryland, College Park in partial fulfillment of the requirements for the degree of Doctor of Philosophy

2018

Advisory Committee:

Professor S. Borağan Aruoba, Co-Chair

Professor John C. Haltiwanger, Co-Chair

Professor Rajshree Agarwal

Professor Serguey Braguinsky

Professor Felipe Saffie

© Copyright by Joonkyu Choi 2018

#### Preface

Who creates jobs in the aggregate economy? A misleading argument that still persists today is that small firms account for the majority of job creation. A now well-known empirical fact in the firm dynamics literature is that young firms, not small firms, are the true engine of job creation (Haltiwanger, Jarmin, and Miranda, 2013). In fact, typical small firms are old and they do not create many jobs. This fact has triggered a great deal of interest among macroeconomists on entrepreneurship, new business formation, and growth and survival dynamics of young firms.

While young firms create massive number of jobs every year, it is important to acknowledge two facts in order to better understand their contribution to the aggregate job creation process. First, jobs created by young firms are not stable. More than 50% of startup firms and the associated jobs are destroyed within their first five years. Second, typical young firms exhibit little or no employment growth (Decker, Haltiwanger, Jarmin, and Miranda, 2014; Hurst and Pugsley, 2011). In fact, a relatively small fraction of young firms grow rapidly and make sustained and large contribution to aggregate job creation. Therefore, it is critical to understand the sources of heterogeneity of young firms, and most importantly, why some young firms grow faster than the others.

Understanding entry, growth, and survival dynamics as well as sources of heterogeneity of young firms has become particularly important in the U.S. The U.S. economy has been experiencing a secular decline in entrepreneurship, business dynamism, and labor market fluidity in the last three decades (Davis and Haltiwanger, 2014; Decker, Haltiwanger, Jarmin, and Miranda, 2014). Moreover, growing

evidence suggests that high-growth young firms are disappearing at a faster pace than non-growing mom and pop stores (Decker, Haltiwanger, Jarmin, and Miranda, 2016).

As part of this research agenda, Chapter 1 of this dissertation provides one answer to the question of "what types of entrepreneurs are more likely to create high-growth young firms, and why?" I show that the entrepreneurs with better fall-back options can afford to take larger business risks, and thus those entrepreneurs are more likely to create high-growth young firms at a cost of higher failure risks. I demonstrate this hypothesis through the lens of a dynamic occupational choice model of entrepreneurship, and test its implications using a comprehensive administrative data that contains 1.7 million U.S. startup firms.

In Chapter 2, I conduct research (jointly with Veronika Penciakova and Felipe Saffie) on job creation effect of the fiscal stimulus during recessions. We focus on the American Recovery and Reinvestment Act (ARRA) enacted during the Great Recession. The primary stated goal of ARRA was promoting aggregate job creation. While most previous studies have focused on identifying the job creation effect of the fiscal stimulus on local regions (local job multiplier), we focus on how firms lobbying behavior affects the allocation of procurement contracts and the magnitude of local job multiplier.

We find that lobbying firms tend to win government contracts with 5% higher probability and 50% larger contracts. We also find that \$ 1 million procurement spending yields on average 11.5 jobs at the Metropolitan Statistical Area (MSA) level, and that effect is entirely driven by procurement spending channeled through

non-lobbying firms. While procurement channeled through lobbying firms has no significant impact on local job creation, \$ 1 million in procurement spending through non-lobbying firms yields 16 jobs. Therefore, our finding cautions that when the allocation of government spending is affected by firms lobbying behavior, job creation effect may be mitigated.

# Dedication

To Jane and Estelle.

#### Acknowledgments

I am deeply indebted to all the people who gave me support, advice, and encouragement during my Ph.D training at the University of Maryland. Completion of this thesis was impossible without them.

I am extremely grateful to my advisors, Borağan Aruoba and John Halti-wanger. John Haltiwanger has introduced me the world of firm dynamics and entrepreneurship, taught me how to tackle macroeconomic questions with large micro-level datasets, and gave me immense support in every step along the process of completing this thesis. Borağan Aruoba has given me extremely helpful and practical advice in every decision I had to make since the very beginning of this work. He taught me how to conduct quantitative and structural research, how to read and write good papers, and how to think like an economist.

I am also indebted to Rajshree Agrawal. She has given me endless trust and support since the first day I met her. She has provided me with numerous opportunities to build my career, but most importantly, taught me the crucial values I need to pursue as a scholar. I am blessed to meet her early in my career. I thank Felipe Saffie for helpful guidance and support during the development of this thesis and for mentoring me as a student and a coauthor. I tremendously benefited from his passionate and kind support; he was always available whenever I needed his advice. I also thank Serguey Braguinsky for much helpful advice with completing my primary research project. His advice was always unique and insightful.

Several others played exceptional roles in making this thesis possible. I thank

John Shea who carefully read every draft of my research and provided me with much detailed comments. His comments were critical inputs in developing this thesis. In addition to my committee, I thank faculty members of the economics department, especially Luminita Stevens, Şebnem Kalemli-Özcan, Nuno Limão and Ina Simonovska for many helpful discussions which undoubtedly improved my research. I also thank Javier Miranda, Emin Dinlersoz, Shawn Klimek, and Nathan Goldshlag at the U.S. Census Bureau for helpful comments and support. I also thank economics department staff members, especially Terry Davis, Vickie Fletcher, Amanda Statland, and Mark Wilkerson for helping me throughout my six years at the University of Maryland.

My former and current colleagues at the University of Maryland have been of immense help in finishing this journey. I thank Jake Blackwood, Hyung Choi, Diyue Guo, Bryan Hardy, Rodrigo Heresi, Lin Hong, Xing Hong, Shen Hui, Karam Jo, Min Kim, Ernest Koh, Jun Hee Kwak, Donggyu Lee, Edith Laget, Tzu-Yao Lin, Heehyun Lim, Hidehiko Matsumoto, Seth Murray, Sungho Noh, Yongjoon Park, Yong Paek, Veronika Penciakova, Svetlana Pivovarova, Cristian Sanchez, Matthew Staiger, Xuezhen Tao, Cody Tuttle, Hsuan Yu, Youngjin Yun, and Jun Wang. I also thank Sangyup Choi, Heechun Kim, Keonwoo Kim, Jongho Park, Kwangyong Park —members of the Junior Scholar Club —for their emotional support and intellectually-enlightening conversations.

My research benefited from support from several institutions. Ewing Marion Kauffman foundation, Ed Snider Center, and the Graduate School of the University of Maryland have provided me with financial support. Because of their support, I

was able to focus on my research without any financial difficulties.

I am eternally grateful to my parents, Kyungha Choi and Seunghee Oh, for their tireless support and love throughout my life. Only with their support I was able to start my journey as an economist.

My daughter Estelle is a source of endless joy and energy of my life. Above all, I thank my wife Jane Kim for her love and endurance throughout the last four years. Nothing would have been possible without her unconditional trust and support. I am deeply indebted to her.

Lastly, I thank God.

### Table of Contents

Preface	)			ii
Dedica	tion			V
Acknov	vledgem	ents		vi
List of	Tables			xi
List of	Figures			xii
	_	urial Risk-	-Taking, Young Firm Dynamics, and Aggregate Impli-	
cati				1
1.1				1
1.2			l of Business Risk Taking	8
1.3			tside Options and Business Risk-Taking	18
	1.3.1		d Measurement	18
	1.3.2	-	ive Statistics	21
	1.3.3		on Analysis	24
		1.3.3.1	Prediction 1: Firm Exit	24
		1.3.3.2	Prediction 2: Growth Dispersion	26
		1.3.3.3	Prediction 3: Growth of Continuers	28
		1.3.3.4	Prediction 4: Interaction with the Hurst-Pugsley Small	
			Business Sector	30
1.4	Gener	al Equilib	rium Model	35
	1.4.1		Description	35
	1.4.2		ry Recursive Competitive Equilibrium	41
1.5	Quant	titative Ar	nalysis	44
	1.5.1	Calibrati	ion	44
	1.5.2	Calibrate	ed Model Properties	50
	1.5.3	Counterf	factual Exercises	54
		1.5.3.1	Removing the Outside Option	54
1.6	Concl	usion		56

2	Lobl	bying, i	Procurement Allocation, and the Employment Effect of Fiscal	
	Stim	nulus (c	oauthored with Veronika Penciakova and Felipe Saffie)	59
	2.1	Introd	uction	59
	2.2	Litera	ture Review	63
	2.3	The A	merican Recovery and Reinvestment Act	66
	2.4	Data 1	Description	68
		2.4.1	Firm-level analysis	68
		2.4.2	MSA-level analysis	71
	2.5	Lobby	ing and ARRA Procurement	72
		2.5.1	Descriptive Statistics	72
		2.5.2	Pooled OLS Specification	75
		2.5.3	Propensity Score Matching Approach	
		2.5.4		82
	2.6	Differe	ential Effects of ARRA Procurement on Employment Growth .	85
		2.6.1	±	85
		2.6.2	Baseline OLS Regressions	
		2.6.3	Instrumental Variables	90
		2.6.4	Regression Results	
	2.7		ple Theoretical Explanation	
	2.8	Conclu	asion	98
	App	endices		101
		A	Appendix for Chapter 1	101
			A.1 Simple Model Solution Derivation	01
			A.2 Robustness Check: Sole-proprietor Sample 1	.05
	Bibl	iograph	v	10

# List of Tables

1.1	Prior earnings and Post earnings	22
1.2	Summary Statistics	23
1.3	Prior Earnings and Initial Performance	24
1.4	Firm Exit Regressions	25
1.5	Growth Dispersion Regressions	28
1.6	Growth Regressions for Continuers	29
1.7	Conditional Employment Growth Regressions	31
1.8	Regression Results with the Hurst-Pugsley Sector Interactions	33
1.9	Externally Calibrated Parameters	45
1.10	Internally Calibrated Parameters	49
1.11	Overall Effect of Removing the Outside Option	55
2.1	Corporate lobbying and procurement allocation (pooled OLS regression)	77
2.2	Pre-matching difference between ARRA lobbying and non-lobbying	' '
2.2	firms	78
2.3	First-stage: post-estimation comparison	80
$\frac{2.5}{2.4}$	Placebo Test	81
2.5	Corporate lobbying and procurement allocation (Second-stage regres-	
	$\operatorname{sion}$ )	82
2.6	First-stage: post-estimation comparison	84
2.7	Compustat Sample: Lobbying and procurement allocation (Second-	85
2.8	stage regression)	
$\frac{2.8}{2.9}$	First-Stage Instrumental Variable Regression Results	93
$\frac{2.9}{2.10}$	<u> </u>	
2.10	MSA Employment Outcomes & ARRA Procurement (IV regression).	
	Compustat Sample MPL and MPK Regression Results)	
A1	Firm Exit Regressions for Sole-Proprietors	
A2	Growth Dispersion Regressions for Sole-Proprietors	
A3	Growth Regressions for Sole-Proprietor Continuers	108
A4	Regression Results with the Hurst-Pugsley Sector Interactions: Sole-	100
	Proprietor Sample	109

# List of Figures

1.1	Value function in period 2	12
1.2	Timing of Events	41
1.3	Average Employment by Firm Age	50
1.4	Life-Cycle Growth and Survival Dynamics in the Model	51
1.5	Experimentation and Innovation of Continuing Firms	53
1.6	Growth of Average Productivity by Firm Age	53
1.7	Growth of Average Productivity by Firm Age with Different Outside	
	Options	54
1.8	Growth of Average Productivity by Firm Age: Benchmark vs. Coun-	
	terfactual	56
2.1	Arra Contracts: Spending by Year (2009-2014)	67
2.2	Distribution of first-year value: ARRA & Non-ARRA contracts (2009-	
	2015)	74
2.3	Distribution of average value of ARRA contracts (2009-2015)	75
2.4	Distribution of ARRA Procurement across MSAs	86
2.5	Distribution of Lobbying share of ARRA Procurement across MSAs .	87
2.6	Distribution of MPL and MPK: Lobbying vs. Non-lobbying firms	97

Chapter 1: Entrepreneurial Risk-Taking, Young Firm Dynamics, and
Aggregate Implications

#### 1.1 Introduction

A long-standing literature in economics, dating back to at least Schumpeter (1942), show that business startups and entrepreneurs play a critical role in innovation, job creation, and productivity growth. Yet, recent studies caution that there is a large heterogeneity in growth dynamics amongst young firms (see, e.g., Decker, Haltiwanger, Jarmin, and Miranda, 2014; Guzman and Stern, 2016; Hurst and Pugsley, 2011; Schoar, 2010). In fact, typical startup firms either exit or exhibit little or no growth, and a small fraction that grow rapidly—so-called high-growth young firms—account for the vast majority of the aggregate contribution of young firms (Decker, Haltiwanger, Jarmin, and Miranda, 2014). However, relatively little is known regarding the economic factors that drive the large differences in the growth dynamics of young firms, and more importantly, the mechanisms through

<sup>&</sup>lt;sup>1</sup>For models of entrepreneurship and innovation, see, among others, Aghion and Howitt (1992), King and Levine (1993), Acemoglu, Akcigit, Bloom, and Kerr (2013), and Acemoglu and Cao (2015). Haltiwanger, Jarmin, and Miranda (2013) and Decker, Haltiwanger, Jarmin, and Miranda (2014) provide empirical evidence on the importance of young firms in aggregate job creation. For evidence on productivity growth, see Haltiwanger, Jarmin, Kulick, and Miranda (2016) and Alon, Berger, Dent, and Pugsley (2017).

which high-growth young firms are created.

In this paper, I propose a key predictor of the early growth trajectories of young firms: the outside options of the business founders. I argue that startup entrepreneurs with higher levels of outside options, which I define by the level of labor income they expect to earn in the event of business failure, are more likely to take larger business risks and thus exhibit a more up-or-out type of firm dynamics. This is because the option to cease business operations and switch to paid employment serves as insurance against business failure, and better insurance enables individuals to take larger risks. Therefore, entrepreneurs with better outside options are more likely to create high-growth young firms at the cost of a higher failure risk, and those with weaker outside options are more likely to create businesses that stay small.<sup>2</sup>

To formalize this argument, I construct a dynamic occupational choice model in which individuals can choose between paid employment and entrepreneurship. I build on earlier work by Vereshchagina and Hopenhayn (2009) and model risk-taking by entrepreneurs as the choice of dispersion in the innovation to their business productivity. I refer to this choice as risky experimentation. Success in experimentation delivers an increase in business productivity, which translates into growth in firm profits and size. Failure in experimentation results in persistent damage to business productivity, which leads to contraction or even to the exit of the firm. Vereshchagina and Hopenhayn (2009) show that the option to return to paid em-

<sup>&</sup>lt;sup>2</sup>I confine the scope of this study to firms that hire at least one employee and exclude nonemployer self-employment activities. Given that the purpose of this study is to examine diverse firm outcomes including employment growth, I consider this as a fair restriction. However, nonemployer businesses are massive in number and deserve investigation as well. For a recent study on the growth outcomes of nonemployer firms, see Fairlie and Miranda (2017).

ployment creates a convexity around the exit margin of the objective functions of the entrepreneurs, which endogenously generates risk-taking incentives. I extend and modify their framework by introducing persistence in the firm productivity process and heterogeneity in labor earnings to generate implications on the relationship between the entrepreneurs' outside options and their post-entry firm dynamics.

I begin the analysis by presenting a stylized two-period version of the model to illustrate the mechanism in its simplest form and to derive analytical solutions that can be mapped into empirically testable predictions. The simple model predicts that firms operated by entrepreneurs with better outside options should exhibit (i) higher exit rates, (ii) more growth dispersion, and (iii) faster growth conditioning on survival compared to firms operated by entrepreneurs with weaker outside options, holding lagged firm productivity constant. I show that it is important to control for lagged productivity to uncover the predicted patterns in the data; the unconditional correlations between outside options and firm exit is ambiguous, given the likely positive correlation between outside options and initial business productivity. The model also implies that when an entrepreneur has strong nonpecuniary incentives for being an entrepreneur (e.g., being one's own boss, having a flexibile work schedule), the impact of his outside option on the three predicted outcomes stated above will be mitigated. This is because if all else is equal, he will be more averse to losing the nonpecuniary benefits of staying in entrepreneurship, and therefore will take fewer risks. This result is consistent with evidence documented by Hurst and Pugsley (2011) that startup business owners who report strong nonpecuniary motives also tend to report a lack of willingness to take risks to grow their firms.

I provide direct empirical evidence for the model's predictions using a panel data of 1.7 million startup firms. To test the model's predictions, one needs a data set that provides information on business founders as well as longitudinal records of their firms, including firm productivity. I construct such unique data set by combining individual- and firm-level administrative databases of the U.S. Census Bureau. It not only contains the demographics and work histories (e.g., earnings, workplaces) of the business founders, but also tracks annually each firm from its first year of operation until exit (if it occurs). Because the outside option is not directly observed in the data, I use the business founders' annual labor earnings prior to business entry as a proxy variable for their outside options. This approach is based on empirical evidence that labor earnings prior to business entry is a strong positive predictor of labor earnings post-business exit, especially for short spells of entrepreneurship.<sup>3</sup> I measure firm productivity by revenue per worker, which is a frequently used measure in the firm dynamics literature.

A major concern of this empirical test is that the outside options of entrepreneurs are likely to be positively correlated with unobserved abilities, such as managerial capabilities, which independently have a positive impact on firm growth and survival outcomes. I find that higher outside options predict higher firm exit rates, once I control for lagged firm productivity and size. A large component of unobserved ability should be captured by lagged firm productivity and size, and if

<sup>&</sup>lt;sup>3</sup>For example, see Williams (2000) and Bruce and Schuetze (2004). I also report a strong positive correlation between prior and post entrepreneurship earnings for the entrepreneurs that exit. The literature also finds that the effect of past entrepreneurship experience on wages is generally smaller than the effect of experience in paid employment. This feature is reflected in the quantitative model in section 1.4.

any effect is left over, it should create a bias toward finding a negative relationship between outside options and firm exit. In addition, it is unclear outside the proposed model mechanism why better unobserved abilities should lead to a larger growth dispersion.

Micro-level theoretical and empirical analyses indicate that the outside options of business founders are important determinants of young firm growth and survival dynamics. Yet, the question remains whether the proposed mechanism have quantitatively meaningful implications for macro-level outcomes such as aggregate output and productivity. To address this question, I embed the stylized model into a heterogeneous agent general equilibrium model and calibrate it to the U.S. economy. I find that a decrease in outside options for startup entrepreneurs can have a sizable impact on aggregate productivity and output. In a counterfactual where the option to return to paid employment is completely removed, aggregate output falls by 8.9%, and aggregate output per worker falls by 4.4%. I find that this result is mainly driven by a reduction in risk-taking by young firms, which results in slower productivity growth along their life cycles. Therefore, outside options are important factors that affect not only young firm growth and survival, but also aggregate output and productivity.

This paper contributes to the entrepreneurship literature that attempts to better understand the gap between a broad population of entrepreneurs with low business growth prospects (e.g., Hamilton, 2000; Hurst and Pugsley, 2011) and a small number of transformative entrepreneurs with strong capabilities and ambition for rapid growth (e.g., Guzman and Stern, 2016; Haltiwanger, Jarmin, Kulick,

and Miranda, 2016). In the developing economy context, Schoar (2010) argues that policy interventions which lack a clear understanding of the difference between those two types of entrepreneurs may result in unintended adverse consequences. While the existing studies tend to adopt such dichotomous view on the types of entrepreneurship, I contribute to this literature by identifying outside options as a relatively continuously distributed source of heterogeneity among entrepreneurs.

This paper also contributes to an emerging literature in firm dynamics and macroeconomics that focuses on the determinants of firm entry and growth along their life cycle. Recent empirical studies found that while young firms make substantial contribution to aggregate job creation and productivity growth, the U.S. economy has been experiencing a secular decline in firm entry rates.<sup>4</sup> In addition, recent studies found that there is a tight linkage between life-cycle dynamics of plants and firms and aggregate productivity (e.g., see Akcigit, Alp, and Peters, 2016; Hsieh and Klenow, 2014). These findings triggered interest among macroeconomists in the life-cycle aspects of firm growth, particularly those of young firms.<sup>5</sup> I contribute to this literature by showing that deterioration in the outside options of entrepreneurs result in a decline in firm entry rates and slower life-cycle growth of young firms.

In addition, this paper contributes to the literature that investigates the experimental aspect of entrepreneurship (e.g., Kerr, Nanda, and Rhodes-Kropf, 2014b). This literature emphasizes that entrepreneurship should be viewed as an experiment

<sup>&</sup>lt;sup>4</sup>For example, see Haltiwanger, Jarmin, and Miranda (2013), Decker, Haltiwanger, Jarmin, and Miranda (2014), Decker, Haltiwanger, Jarmin, and Miranda (2016) and Alon, Berger, Dent, and Pugslev (2017)

<sup>&</sup>lt;sup>5</sup>For recent examples, see Pugsley and Sahin (2015), Arkolakis, Papageorgiou, and Timoshenko (2017), Sedlácek and Sterk (2017), and Moreira (2017).

that can be reversed, and that post-failure options should be taken into consideration in analyzing entrepreneurship decisions. Work by Polkovnichenko (2003) and Vereshchagina and Hopenhayn (2009), and more recently by Manso (2016) and Dillon and Stanton (2017), confirms this idea, and demonstrates that the option to return to paid employment can largely explain why some people enter entrepreneurship despite the low risk premium of entrepreneurs relative to wage earners observed in the cross-sectional earnings distribution (Hamilton, 2000). In a similar vein, empirical studies find that providing insurance against failure from entrepreneurship such as job-protected leave (Gottlieb, Townsend, and Xu, 2016), unemployment insurance (Hombert, Schoar, Sraer, and Thesmar, 2017), and cash transfers (Bianchi and Bobba, 2012) spurs entry to entrepreneurship. I contribute to this literature by providing new empirical evidence such that when lagged firm performance is controlled, higher outside options are associated with higher exit rates and a larger growth dispersion. This result is consistent with the experimentation view of entrepreneurship.

Lastly, this paper is closely tied to existing models of firm dynamics with endogenous innovation choices that involve potential risks (e.g., see, among others, Atkeson and Burstein, 2010, Gabler and Poschke, 2013, Caggese, 2016, and Buera and Fattal-Jaef, 2016). I contribute to this literature by showing that modeling heterogeneity in the post-exit value of firms is important in capturing the dynamics of firms near the entry and exit margins. Typical existing models assume that firms face homogeneous post-failure outcomes by specifying the value of exit as a constant, which is typically set at zero, and focus on other frictions or distortions that affect

firms' innovation decisions. Some of the existing models, such as that of Caggese (2016), recognize that the existence of an exit option generates extra risk-taking incentive for firms, but rarely go further to specify the source of the exit option. I show that modeling the impact of outside options on firm dynamics is important and that using prior earnings can be one way to discipline the distribution of post-exit values. Capturing the dynamics of firms near the entry and exit margins is important, as these firms include startups and young firms, which play an important role in aggregate growth.

The paper is organized as follows. Section 1.2 develops a simple two-period single-agent model that illustrates the risk-taking mechanism. Section 1.3 describes the empirical investigation of the simple model predictions. Section 1.4 extends the simple model to a quantitative heterogeneous agent general equilibrium model, and Section 1.5 describes the model calibration and counterfactual exercises, and Section 1.6 concludes.

### 1.2 A Simple Model of Business Risk Taking

In this section, I present a simple two-period single-agent model of business risk-taking. This model formalizes the mechanism of the hypothesis in its simplest form. It generates a set of predictions on the relationship between outside options and firm growth and survival, which are then empirically tested in Section 1.3. It also serves as a key building block of the quantitative general equilibrium model presented in Section 1.4.

There are two periods, denoted as t = 1, 2. Consider an entrepreneur in t = 1 who is endowed with a business idea  $z_1$  and labor efficiency h.<sup>6</sup> For simplicity, it is assumed that the agent has log utility and all income is comsumed without saving in each period; these assumptions are relaxed later in the quantitative model. In the first period, he hires effective units of labor  $n_1$  and pays  $wn_1$  to workers, where w is the wage per effective unit of labor. He produces output via production function  $y_1 = z_1^{1-\alpha}n_1^{\alpha}$ . His next-period business idea  $z_2$  evolves according to a binomial risky innovation process

$$z_2 = \begin{cases} z_1 e^{\Delta} & \text{with probability} \quad e^{-\gamma \Delta} \\ z_1 e^{-\Delta} & \text{with probability} \quad 1 - e^{-\gamma \Delta} \end{cases}$$

where  $\Delta \geq 0$  is a choice variable and  $\gamma > 0$  is a parameter. Binomial innovation process has been used in existing models of firm dynamics (see, e.g, Atkeson and Burstein, 2010; Buera and Fattal-Jaef, 2016; Caggese, 2016). A key assumption introduced in this model is that  $\Delta$  can be controlled by entrepreneurs, while  $\Delta$  has been treated as a fixed parameter in previous models.<sup>7</sup> This assumption enables the model to predict that some types of firms exhibit larger growth dispersion or higher exit rates than others. Hereafter, I refer to choosing a positive  $\Delta$  as conducting

<sup>&</sup>lt;sup>6</sup>In the simple model, I abstract from entrepreneurship entry decisions and focus on post-entry dynamics. I discuss how outside options affect individuals' entry decisions later in section 1.5.3 using the quantitative model.

<sup>&</sup>lt;sup>7</sup>Atkeson and Burstein (2010) and Buera and Fattal-Jaef (2016) assume that firms can increase the success probability subject to an increasing cost function while  $\Delta$  is a fixed parameter. In Caggese (2016), both success probability and  $\Delta$  are treated as parameters in which  $\Delta$  is specified as risky innovation to fixed cost of operation. Gabler and Poschke (2013) study a firm dynamics model with innovation dispersion choice. An important difference is that firms face a heterogeneous post-exit state in this model, while the post-exit state is assumed to be homogeneous in their model. This feature is at the core of the mechanism described in this paper.

a risky experimentation. The success probability  $e^{-\gamma \Delta}$  is assumed to be decreasing in  $\Delta$ , indicating that riskier experiments are more challenging to implement successfully.

This risky experimentation specification can be thought of as representing several real-world business risk-taking choices made in the post-entry phase. For instance, firms can try to adjust their target customer base: A firm that initially targeted a niche market may try to expand to a broader customer base, in which case it could lose its existing customers in the case of failure. Other examples of risk-taking choices include adding or removing features of a product of service, changing supply-chain systems that may incur disruptions (e.g., Hendricks and Singhal, 2005), and adopting new technologies (e.g., see Holmes, Levine, and Schmitz Jr, 2012 and the references therein).

When the entrepreneur arrives at t=2, he decides whether to stay in business or cease operations after observing the realization of  $z_2$ . In the case of business exit, he switches to paid employment and earns labor income wh, in which case he enjoys the value of  $\ln(wh)$ . If he stays in business, he hires effective units of labor  $n_2$  and earns a profit of  $z_2^{1-\alpha}n_2^{\alpha} - wn_2$ . He chooses  $n_2$  to maximize his utility, so that the value of staying as an entrepreneur in period 2 is

$$\max_{n_2 \ge 0} \ln(z_2^{1-\alpha} n_2^{\alpha} - w n_2) = \ln(\Gamma z_2)$$

where  $n_2^* = \left(\frac{\alpha}{w}\right)^{\frac{1}{1-\alpha}} z_2$  and  $\Gamma = (1-\alpha)\left(\frac{\alpha}{w}\right)^{\frac{1}{1-\alpha}}$ . Therefore, the value function at the

beginning of period 2 can be summarized as

$$V_2(z_2, h) = \max\{\ln(\Gamma z_2), \ln(wh)\}$$

Note that the entrepreneur stays in business if and only if  $z_2 \geq \frac{wh}{\Gamma}$ . Taking  $V_2(z_2, h)$  into account, the entrepreneur in period 1 chooses labor input  $n_1$  and experiment risk  $\Delta$  to maximize expected lifetime utility. Specifically, he solves the problem

$$V_1(z_1, h) = \max_{n_1 \ge 0, \Delta \ge 0} \ln(z_1^{1-\alpha} n_1^{\alpha} - w n_1) + \beta \left\{ e^{-\gamma \Delta} \cdot V_2(z_1 e^{\Delta}, h) + (1 - e^{-\gamma \Delta}) \cdot V_2(z_1 e^{-\Delta}, h) \right\}$$

where  $\beta$  is the time discount factor. The entrepreneur chooses  $n_1$  to maximize period 1 profits, and thus  $n_1^* = (\frac{\alpha}{w})^{\frac{1}{1-\alpha}} z_1$ . The object of interest is the optimal  $\Delta^*$ . It can be shown that for a given  $z_1$ , there exists  $h^*(z_1)$  such that

$$\Delta^* = \begin{cases} \ln(wh) - \ln(\Gamma z_1) + \frac{1}{\gamma} & h \ge h^*(z_1) \\ \bar{\Delta}(\gamma) & 0 \le h < h^*(z_1) \end{cases}$$
(1.1)

where  $\bar{\Delta}(\gamma)$  is a decreasing function of  $\gamma$ . The solution is derived in Appendix A.1. The model predicts that an increase in h leads to a larger  $\Delta^*$ , unless h is too low relative to  $z_1$ .

The core mechanism behind this result is the option value effect. The outside option of switching to paid employment provides a lower bound in the value function of the entrepreneur. This can be seen in Figure 1.1a, in which the value function in t=2 is the upper envelope of the two occupation-specific values  $\ln(wh)$  and

 $\ln(\Gamma z_2)$ . The value function is locally convex around exit threshold  $z_{exit}$  due to the lower bound, and an entrepreneur who has  $z_1$  around this region can increase his expected utility by introducing a risk in  $z_2$ . This endogenous risk-taking behavior is modeled by Vereshchagina and Hopenhayn (2009) in an entrepreneurship context.<sup>8</sup>

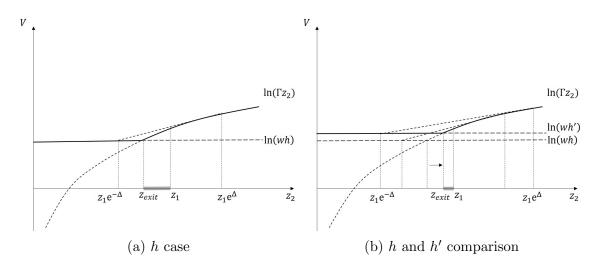


Figure 1.1: Value function in period 2

Notes: Figure (a) shows the value functions of entrepreneurs  $(\ln(\Gamma z))$  and workers  $(\ln(wh))$  in period 2. Figure (b) shows that when there is an increase in labor efficiency from h to h', distance between current productivity and the exit threshold becomes shortened, and risk-taking incentive of the entrepreneur increase.

For a positive  $\Delta$ , the ex post benefit in the case of success is the value gain generated by improving  $z_2$  from  $z_1$  to  $z_1e^{\Delta}$ . The ex post benefit is unbounded above, and the expected marginal benefit with respect to  $\Delta$  diminishes due to the concavity of the utility function and the curvature in the success probability function. In contrast, the ex post cost in the case of failure is bounded below because of the

<sup>&</sup>lt;sup>8</sup>Vereshchagina and Hopenhayn (2009) generates local convexity in the entrepreneurs' value functions along the asset dimension rather than the business productivity dimension. This is done by introducing financial constraints as specified in Evans and Jovanovic (1989) and Holtz-Eakin, Joulfaian, and Rosen (1994). The core mechanism of this paper is not affected by the presence of financial constraints as long as the business productivity follows a persistent process. I introduce financial constraints later in the quantitative model in Section 1.4.

outside option. For any  $z_2$  realizations below  $z_{exit}$ , the entrepreneur will exercise the exit option to minimize the value loss. Hence the lower bound of the ex post cost is determined by the distance between  $z_1$  and  $z_{exit}$ , illustrated by the bold gray line in Figure 1.1a. As shown in Figure 1.1b, an increase in outside option increases the exit threshold  $z_{exit}$ , and shortens the distance between  $z_1$  and  $z_{exit}$  (bold gray line). Therefore an increase in h lowers the expected marginal cost of choosing a large  $\Delta$ , which incentivizes risk-taking behavior.

The positive relationship between h and  $\Delta$  for a given  $z_1$  generates several testable implications on firm growth and survival. First, combining the optimal  $\Delta^*$  and the success probability function  $e^{-\gamma \Delta}$ , the firm exit probability can be derived as

$$Pr(Exit) = \begin{cases} 1 - (\frac{\Gamma z_1}{wh})^{\gamma} e^{-1} & h \ge h^*(z_1) \\ 0 & 0 \le h < h^*(z_1) \end{cases}$$
 (1.2)

Therefore, the model implies that holding  $z_1$  constant, wh and the exit probability in period 2 should be *positively* correlated. Hence, the first prediction is derived.

**Prediction 1.** Controlling for  $z_{t-1}$ , entrepreneurs with higher outside options will exhibit higher firm exit rates in t.

The second prediction is on dispersion of firm growth. Given the specified process for z, a higher  $\Delta^*$  directly implies a larger dispersion in the productivity innovation. In the simple two-period setting, entrepreneurs readily exit when they

 $<sup>^{9}</sup>$ In the quantitative model, I incorporate the direct resource cost of risk-taking. While this cost affects the incentives of individuals for each h, the comparison across different h remains unaltered.

fail in their risk-taking. Thus the observed innovation in z is truncated below, and the model cannot speak to outcomes concerning dispersion. With a straightforward extension to a multiperiod setup, however, it can be shown that even when entrepreneurs stay in business in the case of failure in risk-taking, their  $\Delta_t^*$  are positively associated with h for a given level of  $z_t$ . The intuition is that even though the entrepreneur may not exit in the contemporaneous period after receiving an adverse outcome in  $z_t$ , lower levels of  $z_t$  increase the probability of exiting in the future. Thus outside option wh still affects risk-taking incentives in the same way. Therefore, the model predicts larger dispersion of firm growth for entrepreneurs with higher outside options. This prediction is confirmed in the quantitative model, in which agents are infinitely lived.

**Prediction 2.** Controlling for  $z_{t-1}$ , entrepreneurs with higher outside options will exhibit larger dispersion in growth between t and t-1.

The third prediction is on average firm growth rate conditioning on survival. Because taking a business risk leads to higher probability of exit in case of failure, continuing firms are more likely to consist of risk-taking winners and non-risk-takers. Given that  $z_2 = z_1 e^{\Delta^*}$  for risk-taking winners, their growth rate of z and n is  $e^{\Delta^*} = \frac{wh}{\Gamma z_0} e^{\frac{1}{\gamma}}$ . Since entrepreneurs with higher outside options (wh) tend to take larger risks  $(\Delta^*)$ , they are likely to exhibit faster growth conditioning on survival.

**Prediction 3.** Controlling for  $z_{t-1}$ , entrepreneurs with higher outside options will exhibit faster growth between t-1 and t.

Interaction with Nonpecuniary Motives For Self-employment Empirical studies in the firm dynamics literature indicate that typical startups in the U.S. exhibit little or no growth. The risk-taking mechanism developed in this paper can explain this result if many business founders have low outside options and thus take little or no risks. An alternative hypothesis was put forward by Hurst and Pugsley (2011), who attribute this pattern to the nonpecuniary benefits of self-employment. Using an occupational choice framework, Hurst and Pugsley (2016) show that individuals with strong nonpecuniary motives tend to start businesses in sectors with few scale economies and exhibit no growth. The authors find empirical support for their model using the LBD.

Incorporating their argument into this model yields a unique testable prediction. Following Hurst and Pugsley (2016), the nonpecuniary benefits of entrepreneurship can be incorporated as an additive utility term  $\theta > 0$  in the value function of the entrepreneur. Thus the value function in period 2 can be rewritten as

$$V_2(z_2, h; \theta) = \max\{\ln(\Gamma z_2) + \theta, \ln(wh)\}\$$

The entrepreneur in period 1 then solves the problem

$$\max_{n_1 \ge 0, \Delta \ge 0} \ln(z_1^{1-\alpha} n_1^{\alpha} - w n_1) + \theta + \beta \left\{ e^{-\gamma \Delta} \cdot V_2(z_1 e^{\Delta}, h; \theta) + (1 - e^{-\gamma \Delta}) \cdot V_2(z_1 e^{-\Delta}, h; \theta) \right\}$$

By solving this problem with an strategy identical to the benchmark model, the

<sup>&</sup>lt;sup>10</sup>For example, see Hurst and Pugsley (2011) and Decker, Haltiwanger, Jarmin, and Miranda (2014).

optimal  $\Delta^*$  can be characterized as

$$\Delta^* = \begin{cases} \ln(wh) - \ln(\Gamma z_1) + \frac{1}{\gamma} - \theta & h \ge h^*(z_1; \theta) \\ \bar{\Delta}(\gamma) & 0 \le h < h^*(z_1; \theta) \end{cases}$$
(1.3)

where  $h^*(z_1; \theta)$  is increasing in  $\theta$ . Consequently, exit probability becomes

$$Pr(Exit) = \begin{cases} 1 - (\frac{\Gamma z_1}{wh})^{\gamma} e^{\gamma \theta - 1} & h \ge h^*(z_1; \theta) \\ 0 & 0 \le h < h^*(z_1; \theta) \end{cases}$$
(1.4)

and the growth rate of z and n conditioning on survival is obtained as

$$e^{\Delta^*} = \frac{wh}{\Gamma z_1} e^{\frac{1}{\gamma} - \theta} \tag{1.5}$$

Equations (1.4) and (1.5) indicate that for a given level of h and  $z_1$ , entrepreneurs with strong nonpecuniary motives take fewer risks, and thus exhibit higher survival rates, less growth dispersion, and slower growth conditioning on survival. The intuition behind this result can be understood through Figure 1.1a. Adding  $\theta > 0$  to the entrepreneur's value function is equivalent to a parallel upward shifting of  $\ln(\Gamma z_2)$ , keeping  $\ln(wh)$  constant. In turn, the distance between  $z_1$  and  $z_{exit}$  widens, increasing the expected marginal cost of risk-taking. Intuitively, stronger nonpecuniary motives increase the value cost of closing the business. Therefore, the last prediction of the simple model is obtained.

**Prediction 4.** Controlling for  $z_{t-1}$ , the impact of the outside options on the exit

rate, growth dispersion, and average growth conditioning on survival between t — 1 and t is mitigated for entrepreneurs with strong nonpecuniary motives for self-employment.

Two implications of the simple model should be highlighted. First, it is critical to control for  $z_{t-1}$  to uncover the predicted patterns in the data. As shown below, the data suggest that business founders with higher h tend to enter the market with higher initial values of z.<sup>11</sup> It can be seen from equation (1.1) that when there is a positive correlation between h and z, the unconditional correlation between h and  $\Delta^*$  is ambiguous. Second, the risk-taking incentives generated by the outside option are greatest around the exit margin of z, and diminish as z takes on larger values. Thus, in an environment in which startups enter the market with low levels of z relative to older incumbents, the model predicts that young firms will exhibit larger growth dispersion, higher exit rates, and faster average growth conditioning on survival than older firms. This is a well-known feature of young business dynamics in the U.S (e.g., see Haltiwanger, Jarmin, and Miranda, 2013, Decker, Haltiwanger, Jarmin, and Miranda, 2014).

<sup>&</sup>lt;sup>11</sup>This is partly driven by selection, as individuals tend to enter entrepreneurship when their business ideas are worth pursuing relative to their outside option.

### 1.3 Evidence on Outside Options and Business Risk-Taking

#### 1.3.1 Data and Measurement

To empirically test the predictions established in section 1.2, I combine two administrative databases of the U.S. Census Bureau. The first is the Longitudinal Business Database (LBD), which tracks all U.S. non-farm private establishments and firms with at least one employee since 1976. An establishment is a specific physical location where business activities occur, and all establishments under common operational control are grouped under the same firm ID.<sup>12</sup> The U.S. Census Bureau identifies operational control across business entities through the Economic Censuses and the Company Organization Survey. The LBD tracks business activity information on an annual basis. Data include industry, location, employment, annual payroll, birth, death, and ownership changes (if any) at the establishment level.<sup>13</sup>

Firm growth is measured in four dimensions in this analysis: employment, payroll, revenue, and labor productivity. Payroll and revenue are real annual values, where the CPI-U-RS and the GDP implicit price deflator, respectively, are used for nominal-to-real conversion.<sup>14</sup> Labor productivity is measured by real annual revenue per worker. One limitation of this labor productivity measure is that it does

<sup>&</sup>lt;sup>12</sup>It is important to distinguish establishments from firms, and the Federal Employment Identification Number (EIN) from the firm ID. While most firms start with one establishment and one EIN, high-growth firms often expand to multiple establishments and occasionally obtain multiple EINs.

<sup>&</sup>lt;sup>13</sup>Employment is the number of employees reported to the IRS as of the pay period that includes March 12. For more details on the LBD, see Jarmin and Miranda (2002)

<sup>&</sup>lt;sup>14</sup>For detailed description of the revenue variable in the LBD, see Haltiwanger, Jarmin, Kulick, and Miranda (2016).

not account for cross-industry differences in the contribution of intermediate inputs and prices. Thus, following Haltiwanger, Jarmin, Kulick, and Miranda (2016), all regression analyses use only within-industry variation by including industry by year fixed effects. Industry is classified by the four-digit NAICS code.

The second data is the Longitudinal Employer-Household Dynamics (LEHD), which is a matched employer-employee dataset that covers 95% of private sector jobs. <sup>15</sup> The LEHD combines data from state Unemployment Insurance (UI) earnings records and the Quarterly Census of Employment and Wages (QCEW) of the Bureau of Labor Statistics. It tracks individuals at a quarterly frequency, and provides information on earnings, workplace identifiers, and demographics (e.g., age, race, gender). <sup>16</sup> The highest level of business unit ID in the LEHD is the federal EIN. I integrate federal EIN information from the Business Register with the LBD, and use the crosswalk developed by Haltiwanger, Hyatt, McEntarfer, Sousa, and Tibbets (2014) to merge the LEHD and LBD. <sup>17</sup>

By combining the two datasets, I construct a longitudinal sample of 1.7 million startup firms. This sample is composed of 16 cohorts of startups established between 1999 and 2014 in 31 states.<sup>18</sup> The data contain longitudinally stable identifiers of

<sup>&</sup>lt;sup>15</sup>The LEHD also covers state and local government jobs, but not the federal government jobs. <sup>16</sup>For more details on the LEHD, see Abowd, Stephens, Vilhuber, Andersson, McKinney, Roemer, and Woodcock (2009).

 $<sup>^{17}</sup>$ Due to the different processing of EINs by the IRS and states, some fraction of startup EINs first appear in the two databases in different years. I only include EINs that show consistent startup timing in both datasets: I require that an EIN that belongs to a startup firm ID in the LBD in year t must appear in the LEHD either in the second, third, or fourth quarter of year t-1, or the first quarter of year t.

<sup>&</sup>lt;sup>18</sup>The 31 states are CA, CO, FL, GA, HI, ID, IL, IN, KS, LA, MD, ME, MN, MO, MT, NC, ND, NJ, NM, NV, OR, PA, RI, SC, SD, TN, TX, VA, WA, WI, and WV. Because states joined the LEHD program in a sequential manner, there is a trade-off between state coverage and time length in constructing a balanced sample.

individuals and firms, which enables tracking of individuals' career trajectories, and the formation, growth, and dissolution of the firms they create. For sole proprietorships, founders are identified based on their income tax filings (Schedule C) and EIN applications (Form SS-4). For non-sole-proprietor firms (e.g., corporations), however, business ownership information is not available in the data. Thus, the founders of these firms are approximated with individuals who (1) appear in the initial quarter of business operation, (2) stay within the business for at least three quarters, and (3) are one of the top three earners in the second quarter of operation. This approximation method is a modified version of the approach of Kerr and Kerr (2016), which is frequently adopted in entrepreneurship studies that use the LEHD. Earnings rankings are measured in the second quarter to ensure that the individuals stayed within the firm throughout the quarter, given condition (2). In Appendix A.2, I show that the empirical results are broadly robust to restricting the sample only to sole proprietors.<sup>19</sup>

The key object of interest in this analysis is the outside options of the business founders. In the simple model in section 1.2, outside options are defined by the labor income founders expect to earn in the case of firm exit. In the empirical analysis, founders' outside option is proxied by their real annual labor earnings prior to business entry.<sup>20</sup> This is based on the assumption that a founder's expected income after his business failure would be in a range similar to his prior earnings. While some empirical studies in the literature supports this assumption (e.g., Bruce

<sup>&</sup>lt;sup>19</sup>In future drafts, additional robustness checks on founder identification will be conducted.

<sup>&</sup>lt;sup>20</sup>Specifically, I measure prior earnings by the sum of real quarterly earnings from the most recent four consecutive full-quarter main jobs. The CPI-U-RS is used for nominal-to-real conversion.

and Schuetze, 2004; Williams, 2000), I also report a simple OLS regression of post-entrepreneurship earnings on prior earnings in Table 1.1, in which both objects are normalized by economy-wide average real earnings to remove the aggregate time trend. I find a statistically significant coefficient of 0.52, and that prior earnings alone account for 30% of the variation in post-entrepreneurship earnings.

Businesses are not included in the analysis if their founders did not have any prior earnings records. However, I find that average and median founder ages are around 39, and that most founders have at least one year of prior earnings record. Therefore, I conclude that such sample selection is not likely to bias the results. I further restrict the sample to the businesses where the average founder ages are between 25 and 54. The upper bound is imposed to stay reasonably far from any retirement considerations and the lower bound is set to increase the probability of capturing the young founders' full-year and full-time jobs prior to business entry. The empirical results are robust to the sample without the age restrictions.

### 1.3.2 Descriptive Statistics

Table 1.2 provides summary statistics of the sample. Panel A shows the employment and productivity distribution of startups and employment growth rates of year-to-year continuers up to age five. The employment growth rate between year t-1 and t is calculated as  $\frac{Emp_t-Emp_{t-1}}{(Emp_t+Emp_{t-1})/2}$ . This measure is known as the Davis-Haltiwanger-Schuh (DHS) growth rate (Davis, Haltiwanger, and Schuh, 1996), and

Table 1.1: Prior earnings and Post earnings

	Log (relative) Post Earnings
Log (relative) Prior Earnings	0.520***
	(0.001)
Constant	-0.115***
	(0.001)
Obs.	1090000
R-sq	0.300

Notes: The table reports results for OLS regressions in which the independent variable is founder log prior earnings and the dependent variable is the same individual's post-entrepreneurship earnings. Observation counts are rounded to the nearest 10,000 to avoid potential unwarranted disclosure. Prior earnings are measured as the sum of real quarterly earnings at the most recent four consecutive full-quarter main jobs prior to business entry. Likewise, post earnings are measured as the sum of real quarterly earnings at the first four consecutive full-quarter jobs post business exit. Both earnings are normalized by economy-wide average real earnings obtained from the LEHD QWI. \*\*\*, \*\*, and \* indicate significance at the 1%, 5%, and 10% level, respectively.

is standard in the firm dynamics literature.<sup>21</sup> The employment growth rate distribution is weighted by employment. Weighting the distribution by employment, together with the use of the DHS growth measure, minimizes the negative relationship between size and growth generated by scale differences. Panel A in Table 1.2 reconfirms the previous findings in the literature. Most startup firms are small, and typical continuing young firms exhibit only little growth; the median growth rate is 1%. However, young firm growth rates show large dispersion and positive skewness, driving the mean up to 6.5%.

The relative labor productivity of each startup is measured as the deviation from its own industry's average labor productivity in the same year. Reported labor productivity statistics are calculated from an (unweighted) distribution that

<sup>&</sup>lt;sup>21</sup>The DHS growth measure mitigates the problem known as the "regression to the mean effect," and it is symmetric around zero. It is identical to the log differences up to a second-order approximation. For details, see Törnqvist, Vartia, and Vartia (1985), Davis, Haltiwanger, and Schuh (1996), and Haltiwanger, Jarmin, and Miranda (2013).

Table 1.2: Summary Statistics

	Obs.	Mean	Median	St. Dev.	p90	p10	
Par	nel A. F	irm Dis	tributions	}			
Age 0 employment	$1.7 \mathrm{m}$	6.7	3	21.4	14	1	
Age 0 labor productivity	$1.7 \mathrm{m}$	049	024	1.07	1.18	-1.24	
Age 1-5 emp. growth	4.4m	.065	.01	.45	.56	38	
Panel B. Founder Distributions							
Founder log prior earnings	1.7m	10.6	10.6	0.83	11.61	9.62	
Founder age	1.7m	38.9	38.7	7.48	49.5	28.7	

Notes: The table reports summary statistics for the data. To avoid potential unwarranted disclosure, median, 90th, and 10th percentiles are calculated as the averages of their one-percentile neighborhood (e.g., p90 = (p89 + p91)/2). Relative labor productivity for each startup is measured by the deviation from its industry average labor productivity in its startup year, and the statistics are calculated from the (unweighted) distribution that combines observations between 1999 and 2013 across all industries. Prior earnings of each founder are measured by the sum of real quarterly earnings at four consecutive full-quarter main jobs prior to the startup quarter. Real quarterly earnings are evaluated in 2012 Q1 dollars using the CPI-U-RS. If more than one founder is identified for a given business, averages are taken across founders.

combines all observations between 1999 and 2013 across all industries. Labor productivity for the average startup firm is 4.9% lower than its industry's average. This estimate is in line with Foster, Haltiwanger, and Krizan (2001), who find that entering plants in the U.S. manufacturing sector in 1987 had 7% to 8% lower labor productivity than incumbent plants. As stated above, average and median founder ages are around 39, indicating that typical startup founders are at the peak of their prime working age. Converting logs to levels, the average founder prior earnings is around \$40,134 in 2012 dollars.

Table 1.3 shows the relationship between founder prior earnings and their initial performance, viewed through a simple OLS regression. Age 0 employment

and relative labor productivity are regressed on founder log prior earnings. Industry by year fixed effects are included for the reasons explained above. Unsurprisingly, founders with higher prior earnings tend to start businesses with more employees and higher productivity levels. Estimated coefficients indicate that a one standard deviation increase in log prior earnings is associated with 1.93 more employees and 12.5% higher labor productivity in the firm's first year of operation.

Table 1.3: Prior Earnings and Initial Performance

	Age 0 emp.	Age 0 relative labor productivity
Founder log prior earnings	2.337***	0.151***
	(0.343)	(0.0108)
Ind-Year FE	Yes	Yes
Obs.	$1.7\mathrm{m}$	$1.7\mathrm{m}$
R-sq	0.06	0.21

**Notes:** The table reports results for OLS regressions in which the independent variable is founder log prior earnings. Industry by year fixed effects are included in both regressions, where industry is defined by the four-digit NAICS code. Standard errors are clustered at the industry (NAICS4) level. \*\*\*, \*\*, and \* indicate significance at the 1%, 5%, and 10% level, respectively.

# 1.3.3 Regression Analysis

#### 1.3.3.1 Prediction 1: Firm Exit

Do higher outside options of business founders predict higher firm exit rates? To answer this question, I estimate linear probability regression models in which the dependent variable is the firm exit indicator.<sup>22</sup> Results are reported in Table 1.4. The first column shows the simplest case in which prior earnings is the only independent variable. The estimated coefficient indicates that the unconditional

<sup>&</sup>lt;sup>22</sup>Results are robust to using logit or probit regressions.

correlation between prior earnings and firm exit is negative. As explained in Section 1.2, however, it is critical to control for lagged business productivity to uncover the patterns predicted by the model. This is particularly important as the findings reported in Table 1.3 combined with Equation (1.1) imply an ambiguous unconditional correlation between prior earnings and exit rates.

Table 1.4: Firm Exit Regressions

Dependent Variable: Firm exit indicator								
	(1)	(2)	(3)	(4)				
Log prior earnings	-0.005**	0.003***	0.019***	0.019***				
	(0.002)	(0.001)	(0.001)	(0.001)				
Lagged log employment		-0.039***	-0.055***	-0.055***				
		(0.001)	(0.002)	(0.002)				
Lagged log labor prod.			-0.082*** (0.005)	-0.083*** (0.005)				
Lagged log wage				-0.001 (0.003)				
Founder average age				-0.001*** (0.000)				
Founder male share				0.004** (0.002)				
Ind-Year FE	No	Yes	Yes	Yes				
Firm age FE	No	Yes	Yes	Yes				
State FE	No	No	No	Yes				
Obs. (firm-year)	4920000	4920000	4920000	4920000				
R-sq	0.00	0.03	0.09	0.10				

**Notes:** The table reports results for a linear probability regression in which the dependent variable is firm exit indicator. Observation counts are rounded to the nearest 10,000 to avoid potential unwarranted disclosure. Standard errors are clustered at the industry (NAICS4) level. \*\*\*, \*\*, and \* indicate significance at the 1%, 5%, and 10% level, respectively.

Indeed, regression results support the model's prediction when lagged business productivity indicators are included in the regression. Column (2) shows results with

lagged firm size and firm age controls, as well as industry by year fixed effects. Firm size is known to be highly correlated with firm productivity, and including it turns the coefficient on prior earnings from negative to positive. When labor productivity is also included in the regression as in column (3), the positive relationship between prior earnings and firm exit probability is strengthened by an order of magnitude. Column (4) shows that this result is robust to including additional controls. First, I include the lagged log firm wage as an additional indicator of firm productivity, where the firm wage is calculated as the payroll per worker in the first quarter of each year. Second, I control for the average age and fraction of males among founders within each firm to account for individual characteristics that might be correlated with both prior earnings and their risk preferences. Older individuals may be more risk-averse, and males are known to be more prone to risk-taking (Laasch and Conaway, 2009). To the extent that age is positively associated with prior earnings and that gender wage gaps exist, omitting age and gender will introduce a downward and upward bias, respectively, in the estimate of the coefficient on prior earnings. Lastly, state fixed effects are included, as some states have a more dynamic business environments than others, and business dynamism is known to be positively associated with labor market conditions (Davis and Haltiwanger, 2014).

## 1.3.3.2 Prediction 2: Growth Dispersion

Do business founders with better outside options exhibit larger firm growth dispersion? To answer this question, a firm-level growth dispersion measure is con-

structed following Castro, Clementi, and MacDonald (2009). To begin with, I compute the portion of firm growth that is systematically predicted by firm-level and aggregate factors by estimating the following regression model:

$$\Delta \ln Y_{ijt} = \beta_0 + \beta_1 \ln(emp)_{ijt-1} + Firmage_{ijt} + \eta_{jt} + \alpha_i + \epsilon_{ijt}$$
 (1.6)

where  $\Delta \ln Y_{ijt}$  is the DHS growth rate of either log revenue or log labor productivity,  $Firmage_{ijt}$  is a series of dummies for firm age,  $\eta_{jt}$  are industry by year fixed effects, and  $\alpha_i$  is a firm fixed effect. Only year-to-year continuers are included in the estimation. The object of interest is  $\hat{\epsilon}_{ijt} = \Delta \ln Y_{ijt} - \widehat{\Delta \ln Y_{ijt}}$ , the deviation of growth from its conditional mean. Larger firm-level growth dispersion corresponds to larger squared deviations from its conditional mean,  $\hat{\epsilon}_{ijt}^2$ . It is assumed that  $\hat{\epsilon}_{ijt}^2 = f(X_{ijt}) + \nu_{ijt}$ , where  $X_{ijt}$  is a vector of factors that are systematically related to firm-level growth dispersion. Approximating  $f(\cdot)$  linearly, I estimate regression equation (1.7) to test the model prediction in which  $\tilde{X}_{ijt}$  is a vector of factors other than log prior earnings. Table 1.5 shows the results.

$$\epsilon_{ijt}^2 = \beta_0 + \beta_1 \log \text{ prior earnings}_i + \beta_2 \tilde{X}_{ijt} + \nu_{ijt}$$
 (1.7)

Results are consistent with the model prediction: Higher log prior earnings predict larger firm-level growth dispersion.  $\epsilon_{ijt}^2$  ( $\Delta$  Rev) and  $\epsilon_{ijt}^2$  ( $\Delta$  Prod) are the squared deviations obtained from estimating Equation (1.6), where  $Y_{ijt}$  are revenue

<sup>&</sup>lt;sup>23</sup>The results are robust to using absolute deviations  $|\hat{\epsilon}_{ijt}|$ .

Table 1.5: Growth Dispersion Regressions

	$\epsilon^2(\Delta \text{ Rev})$	$\epsilon^2(\Delta \text{ Rev})$	$\epsilon^2(\Delta \text{ Prod})$	$\epsilon^2(\Delta \text{ Prod})$
Log prior earnings	0.014***	0.007***	0.009***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)
Lagged log employment		-0.003*** (0.000)		-0.011*** (0.000)
Ind-Year FE	No	Yes	No	Yes
Firm age FE	No	Yes	No	Yes
Obs. (firm-year)	4410000	4410000	4410000	4410000
R-sq	0.00	0.03	0.00	0.02

**Notes:** The table reports results from estimating Equation (1.7).  $\epsilon_{ijt}^2$  ( $\Delta$  Rev) and  $\epsilon_{ijt}^2$  ( $\Delta$  Prod) are the squared deviations obtained from Equation (1.6), where  $Y_{ijt}$  are revenue and labor productivity, respectively. Observation counts are rounded to the nearest 10,000 to avoid potential unwarranted disclosure. \*\*\*, \*\*, and \* indicate significance at the 1%, 5%, and 10% level, respectively.

and labor productivity, respectively. Two sets of regression models are estimated for robustness analysis. In the first and third columns, only prior earnings is used as the independent variable. The second and fourth columns add controls for firm size, age, and industry by year fixed effects. Controlling for firm age accounts for other mechanisms that can explain the high growth dispersion of young firms, such as the learning and selection effects pioneered by Jovanovic (1982). Industry by year fixed effects are included to control for time-varying industry-level factors such as uncertainty shocks (see, e.g., Bloom, 2009) that can affect firm-level idiosyncratic growth dispersion.

#### 1.3.3.3 Prediction 3: Growth of Continuers

The simple model also predicts that conditioning on survival, entrepreneurs with high outside options exhibit faster firm growth. To test this prediction, firm

growth is regressed on the full set of controls used in Section 1.3.3.1. Firm growth is measured in four dimensions: revenue, labor productivity, payroll, and employment. All growth rates are measured using the DHS method. Results are reported in Table 1.6.

Table 1.6: Growth Regressions for Continuers

	Δ Revenue	$\Delta$ Prod.	$\Delta$ Payroll	$\Delta$ Emp.
Log prior earnings	0.014***	0.02***	0.01***	-0.006**
	(0.002)	(0.003)	(0.002)	(0.002)
Lagged log employment	-0.013***	$0.048^{***}$	$0.007^{***}$	-0.063***
	(0.003)	(0.003)	(0.002)	(0.004)
Lagged log labor prod.	-0.05***	-0.168***	0.098***	0.117***
	(0.005)	(0.005)	(0.004)	(0.006)
Lagged log wage	0.014***	-0.002	-0.104***	0.019***
	(0.003)	(0.003)	(0.003)	(0.005)
Founder average age	-0.002***	-0.000**	-0.001***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
Founder male share	0.013***	0.032***	0.007**	-0.019***
	(0.003)	(0.005)	(0.003)	(0.005)
Ind-Year FE	Yes	Yes	Yes	Yes
Firm age FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Obs.	4410000	4410000	4410000	4410000
R-sq	0.05	0.11	0.06	0.12

**Notes:** The table reports results for OLS regression of firm growth on prior earnings. All growth measures are calculated as the DHS growth rate. Observation counts are rounded to the nearest 10,000 to avoid potential unwarranted disclosure. Standard errors are clustered at the industry (NAICS4) level. \*\*\*, \*\*, and \* indicate significance at the 1%, 5%, and 10% level, respectively.

Results are consistent with the model's prediction, with the exception of employment growth. A one standard deviation increase in log prior earnings is associated with a 1.1% annual increase in revenue, a 1.6% increase in labor productivity,

a 0.8% increase in payroll, and a 0.5% decline in employment on average. Note that these are estimates only of the direct effects of prior earnings in each year. The negative average effect of log prior earnings on employment growth can be explained by the larger growth dispersion of founders with higher prior earnings. That is, founders with higher prior earnings must exhibit faster growth conditioning on expanding, and also faster decline conditioning on contracting. If the latter effect dominates the former, one may find the impact of prior earnings to be negative on average. To see whether this explanation is supported by the data, employment growth regressions are re-estimated separately for firms with positive employment growth (expansions) and negative employment growth (contractions). Results are shown in Table 1.7, in which other control variables are suppressed for simple exposition. Estimation results are consistent with the explanation, indicating that employment growth patterns are also consistent with model predictions.

# 1.3.3.4 Prediction 4: Interaction with the Hurst-Pugsley Small Business Sector

Lastly, the model predicts that all results presented so far will be mitigated for business founders with a strong preference for self-employment. Although this preference is not observable directly, Hurst and Pugsley (2016) show that individuals with strong nonpecuniary motives are likely to be concentrated in sectors with small natural scale. Their intuition is that if the primary goal is to become a business owner and not to earn large profits, those individuals will do so in the most cost-

Table 1.7: Conditional Employment Growth Regressions

	$\Delta \text{ Emp} > 0$	$\Delta \text{ Emp} < 0$
Log prior earnings	0.016***	-0.023***
	(0.002)	(0.002)
Lagged log labor prod.	0.06***	0.049***
	(0.004)	(0.002)
Lagged log employment	-0.164***	0.095***
	(0.003)	(0.003)
Ind-Year FE	Yes	Yes
Firm age FE	Yes	Yes
Full controls	Yes	Yes
Obs.	1520000	1190000
R-sq	0.38	0.17

**Notes:** The table reports results for the OLS regression in which the dependent variable is annual employment growth rate, conditioning on employment expansions and contractions separately. Observation counts are rounded to the nearest 10,000 to avoid potential unwarranted disclosure. Standard errors are clustered at the industry (NAICS4) level. \*\*\*, \*\*, and \* indicates significance at the 1%, 5%, and 10% levels, respectively.

effective way. In their model, differences in natural scale are driven by heterogeneous fixed costs of operation; hence, small natural scale sectors are the least costly to enter. Hereafter, such small natural scale sectors are labeled collectively as the Hurst-Pugsley (HP) sector.

Following Hurst and Pugsley (2016), the HP sector is defined by the top 40 (out of 294) four-digit NAICS industries in terms of small business intensity. Small business intensity of industry j,  $x_j$ , is calculated by

$$x_j = \frac{s_j}{\sum_k s_k}$$

where  $s_j$  is the number of small businesses (fewer than 20 employees) in industry j. The denominator is the sum of  $s_k$  across all industries. Then an indicator variable  $\mathrm{HP}_j$  is created, which takes a value of one if industry j is in the HP sector and zero otherwise. The regression models presented above are re-estimated including the interaction term between the HP indicator and log prior earnings. Results for key dependent variables are presented in Table 1.8. For all regression models, the estimated coefficients on HP interaction terms have the opposite signs as the coefficients on log prior earnings, and in many cases are statistically significant. Therefore, the results are consistent with the model prediction.

Table 1.8: Regression Results with the Hurst-Pugsley Sector Interactions

	(1) Fxit	$\begin{array}{c} (2) \\ \wedge \text{Rev} \end{array}$	$\begin{array}{c} (3) \\ A \text{ Proh} \end{array}$	$\frac{(4)}{\Lambda \text{ Fmp}} > 0$	$\begin{array}{c} (5) \\ \wedge \text{ Fmp } < 0 \end{array}$	$\begin{array}{c} (6) \\ (A \text{ Prod}) \end{array}$
Log prior earnings	0.023***		0.027***	0.018*** (0.003)	-0.027*** (0.002)	0.008***
$\mathrm{HP} \times \log \mathrm{prior} \; \mathrm{earnings}$	$-0.006^{***}$ (0.001)	$-0.011^{***}$ (0.003)	$-0.01^{**}$ (0.005)	-0.003 $(0.005)$	0.006 $(0.004)$	$-0.002^{***}$ (0.000)
Lagged log labor prod.	$-0.083^{***}$ (0.005)	-0.059*** (0.005)	$-0.168^{***}$ (0.005)	$0.06^{***}$ (0.004)	$0.049^{***}$ $(0.002)$	
Lagged log employment	-0.055*** (0.002)	$-0.013^{***}$ (0.003)	0.047***	$-0.164^{***}$ (0.003)	0.095***	-0.003***
Ind-Year FE Firm age FE	Yes	Yes Yes	Yes Yes	Yes	Yes	Yes
Full controls Obs. R-sq	$\begin{array}{c} \text{Yes} \\ 4920000 \\ 0.10 \end{array}$	Yes 4410000 0.05	Yes 4410000 $0.11$	$\begin{array}{c} \mathrm{Yes} \\ 1520000 \\ 0.38 \end{array}$	$\begin{array}{c} \mathrm{Yes} \\ 1190000 \\ 0.17 \end{array}$	0.05 $0.03$

Observation counts are rounded to the nearest 10,000 to avoid potential unwarranted disclosure. Standard errors are clustered at the industry level Notes: The table reports results for linear regressions re-estimated after including the interaction between HP indicator and log prior earnings. (NAICS4). \*\*\*, \*\*, and \* indicate significance at the 1%, 5%, and 10% level, respectively.

The Need for a Quantitative General Equilibrium Model The empirical evidence mostly supports the prediction from the simple model that if business founders have better outside options, they tend to take larger business risks and thus exhibit a more up-or-out type of firm dynamics. This result suggests that factors affecting the outside options of entrepreneurship (e.g. labor market frictions) would alter risk-taking behavior by startups and growth dynamics along their life cycle. However, whether this channel can be translated into a quantitatively significant effect on the aggregate economy is unclear. For instance, because the share of business activity accounted for by young firms is small, changes in their risk-taking behavior may not generate sufficiently strong forces to affect macro-level outcomes.

Therefore, one needs a structural macroeconomic model disciplined by data to investigate the quantitative importance of outside options and young firms' risk-taking behavior in shaping aggregate outcomes. Structural macro models also provide a useful laboratory to conduct experiments that cannot be done otherwise, such as altering the outside options of startup entrepreneurs while holding all other factors constant. In the following section, I embed the simple model presented in Section 1.2 into a heterogeneous agent general equilibrium model to conduct this analysis.

## 1.4 General Equilibrium Model

### 1.4.1 Model Description

Environment There is a continuum of individuals and time is discrete. At the beginning of each period t, individuals randomly die with probability  $\zeta$ , and the same mass of individuals newly enter the economy. Individuals live infinitely unless they are hit by the death shock. Upon entry to the economy, individuals receive assets  $a_0$  from a distribution  $\mu_a(a)$ , and draw their effective units of labor h from a Pareto distribution F(h),

$$F(h) = 1 - h^{-\lambda}, \qquad h \ge 1$$
 (1.8)

Also, individuals draw their initial business productivity z from a log-normal distribution where

$$\ln(z) \sim \mathcal{N}(\mu_z(h), \sigma_z^2) \tag{1.9}$$

It is assumed that

$$\mu_z(h) = \underline{z} + \rho h \tag{1.10}$$

where  $\underline{z}$  is a parameter that governs the overall location of the z distribution. The dependence of  $\mu_z(h)$  on h reflects the notion that individuals with different levels of effective units of labor may have access to different types of business ideas. Individuals also draw preference for entrepreneurship  $\theta \in \{0, \bar{\theta}\}$ , where  $\theta = \bar{\theta} > 0$  with probability  $p_{\theta}$  and 0 with probability  $1 - p_{\theta}$ . For simplicity,  $\theta$  is assumed to stay

constant throughout individuals' lifetime.

Occupational Choice At the beginning of each period, individuals decide whether to become a worker or an entrepreneur. When individuals choose to become paid workers, they supply their effective units of labor h to entrepreneurs and earn wages wh, where w is the wage rate per effective unit of labor. Workers receive interest payments by depositing their assets a, and consume c out of (1 + r)a + wh. Thus, the workers' problem is defined by

$$V^{W}(a, z, h, \theta) = \max_{c \ge 0} u(c) + \beta(1 - \zeta)V(a', z, h, \theta)$$
 (1.11)  
s.t  $a' = (1 + r)a + wh - c \ge 0$ 

For simplicity, h is assumed to stay constant if the individuals stay in paid employment. Hereafter, next-period variables are denoted with superscript '. The value function V is defined as

$$V(a, z, h, \theta) = \max\{V^{W}(a, z, h, \theta), V^{E}(a - \phi, z, h, \theta)\}$$
(1.12)

where  $V^E(a,z,h,\theta)$  is the value of becoming an entrepreneur and  $\phi$  is a fixed entry and exit cost. When individuals enter entrepreneurship, they employ effective units of labor n and rent physical capital k to produce output y via production function  $y = z^{1-\nu}(k^{\alpha}n^{1-\alpha})^{\nu}$ .  $\alpha \in (0,1)$  is the capital production share, and  $\nu \in (0,1)$  is the decreasing returns to scale parameter, which stems from the limited span of control

as in Lucas (1978).

Risky Experimentation and Incremental Innovation A key feature of the model is that entrepreneurs can attempt to enhance their productivity levels through two means of innovation: risky experimentation or incremental innovation. When an entrepreneur engages in risky experimentation, his next-period business productivity z' evolves according to a binomial risky process, as in the simple model in Section 1.2:

$$z' = \begin{cases} ze^{\Delta^R} & \text{w/ prob.} \quad e^{-\gamma \Delta^R} \\ ze^{-\Delta^R} & \text{w/ prob.} \quad 1 - e^{-\gamma \Delta^R} \end{cases}$$

That is, his business productivity either increases or decreases by  $\Delta^R$  percent.  $\gamma > 0$  is the success probability elasticity of risk choice, and  $\Delta^R$  is assumed to be a choice variable implying that entrepreneurs can decide how much risk to take. To conduct risky experiment, entrepreneurs must pay an experimentation cost Fz > 0.

When an entrepreneur attempts to achieve incremental innovation, his nextperiod business productivity z' evolves as

$$z' = \begin{cases} ze^{\Delta^I} & \text{w/ prob.} \quad u \\ z & \text{w/ prob.} \quad 1 - u \end{cases}$$

where the ex post outcome is bounded below by the status quo. Innovation step size  $\Delta^I > 0$  is a fixed parameter, and success probability u is a choice variable.

Entrepreneurs can increase u subject to the cost function

$$R_I = \chi z u^{\psi}$$

where  $\frac{1}{\psi} < 1$  is the success probability elasticity of research cost  $R_I$ , and  $\chi > 0$  is a scaling parameter. This is a standard functional specification for R&D activity in the endogenous growth and innovation literature.<sup>24</sup>

Incremental innovation is introduced in the quantitative model to avoid potential bias in quantitative evaluation of the risky experimentation channel. In the absence of incremental innovation, the firm productivity distribution in the model is completely pinned down by the initial z distribution and the risky experimentation performed by young firms whose z's are not too far from the exit margin. In reality, however, large, mature firms often engage in innovation activities, including R&D. Therefore, the quantitative relevance of the risky experimentation channel will be overemphasized if aggregate research statistics are targeted in calibration and the innovation tools used by mature incumbent firms are not introduced in the model.

Financial Market There are financial intermediaries who own a technology that transforms consumption goods into physical capital, and vice versa, at a one-to-one rate. Financial intermediaries receive deposits from workers and entrepreneurs, and use the transformation technology to rent physical capital to entrepreneurs. It is assumed that the financial market is perfectly competitive, and thus the capital rental

<sup>&</sup>lt;sup>24</sup>For example, see Acemoglu, Akcigit, Bloom, and Kerr (2013) and Akcigit and Kerr (forthcoming).

rate is  $r + \delta$ , where  $\delta > 0$  is the capital depreciation rate. Financial intermediaries require collateral from entrepreneurs when engaging in capital rental contracts. Denoting  $m \geq 0$  as the collateral amount, entrepreneurs can borrow capital only up to a multiple of their collateral, i.e.,  $k \leq \lambda m$ .<sup>25</sup> At the beginning of each period, entrepreneurs decide how much to consume out of their assets a, and the remainder m = a - c is put up as collateral. Financial intermediaries pay interest on the collateral at the beginning of the following period.

Entrepreneurs' Optimization Problem The optimization problem for an entrepreneur who engages in risky experimentation can be expressed as

$$V^{E,R}(a,z,h,\theta) = \max_{c,n,k,\Delta^R} u(c) + \theta + \beta(1-\zeta) \left[ e^{-\gamma \Delta^R} \tilde{V}(a',ze^{\Delta^R},h,\theta) + (1-e^{-\gamma \Delta^R}) \tilde{V}(a',e^{-\Delta^R},h,\theta) \right]$$

$$(1.13)$$

s.t 
$$a' = m + z^{1-\nu} (k^{\alpha} n^{1-\alpha})^{\nu} - wn - (r+\delta)k - Fz \cdot \mathbb{I}(\Delta^R > 0) \ge 0$$
  

$$m = (1+r)a - c \ge 0$$

$$k < \lambda m$$

where

$$\tilde{V}(a', z', h, \theta) = \max \{ V^{E}(a', z', (1 - \delta_h)h, \theta), V^{W}(a' - \phi, z', h, \theta) \}$$
(1.14)

 $<sup>^{25}</sup>$ This simple financial constraint is widely used in this class of models due to its tractability. For example, see Buera and Shin (2013) and Moll (2014).

and  $\mathbb{I}(\Delta^R > 0)$  is an indicator that takes a value of one if  $\Delta^R > 0$ . I assume that labor efficiency depreciates by  $\delta_h$  percent if an individual stays in entrepreneurship. This assumption reflects the idea that individuals who enter entrepreneurship and then return to paid employment generally end up on a worse career trajectory.<sup>26</sup>

Likewise, the optimization problem for entrepreneurs who attempt to achieve incremental innovation can be expressed as

$$V^{E,I}(a, z, h, \theta) = \max_{c, n, k, u} u(c) + \theta + \beta (1 - \zeta) \left[ u \tilde{V}(a', z e^{\Delta I}, h, \theta) + (1 - u) \tilde{V}(a', z, h, \theta) \right]$$
(1.15)

s.t 
$$a' = m + z^{1-\nu} (k^{\alpha} n^{1-\alpha})^{\nu} - wn - (r+\delta)k - \chi z u^{\psi} \ge 0$$
  

$$m = (1+r)a - c \ge 0$$
  

$$k < \lambda m$$

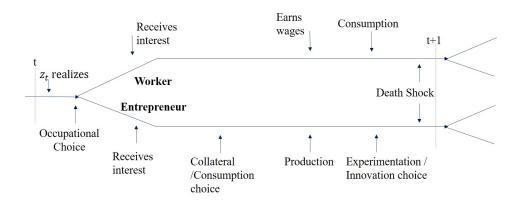
Entrepreneurs in a given period decide between risky experimentation or incremental innovation (or no innovation), and cannot conduct both. Hence, the value function of being an entrepreneur is

$$V^{E}(a, z, h, \theta) = \max\{V^{E, R}(a, z, h, \theta), V^{E, I}(a, z, h, \theta)\}$$
(1.16)

Figure 1.2 summarizes the timing of events.

 $<sup>^{26} \</sup>mathrm{For}$  example, see Williams (2000), Bruce and Schuetze (2004), and Baptista, Lima, and Preto (2012).

Figure 1.2: Timing of Events



## 1.4.2 Stationary Recursive Competitive Equilibrium

In the quantitative analysis, I focus on a stationary recursive competitive equilibrium. The interest rate r is treated as exogenously given, so that the model can be considered to be a small open economy. The state variables of individuals are assets, effective labor, business productivity, preference for entrepreneurship, and occupation.  $\mathbf{A} = [0, \infty)$  is the set of possible asset holdings a, and  $\mathbf{Z} = [0, \infty)$  is the space of business productivity z.<sup>27</sup> Effective labor h is defined over  $\mathbf{H} = [0, \infty)$ , and preference for entrepreneurship  $\theta$  is defined over  $\mathbf{\Theta} = \{0, \overline{\theta}\}$  as above. Occupation is defined as  $o \in \mathbf{O} = \{w, er, ei\}$ , where each element in  $\mathbf{O}$  represents being a worker, an entrepreneur conducting risky experimentation, and an entrepreneur conducting incremental innovation, respectively. Then the distribution of individuals  $\mu$  is defined as a probability measure  $\mu(a, z, h, \theta, o) : B \to [0, 1]$ , where B is the Borel  $\sigma$ -algebra generated by the open sets of the product space

<sup>&</sup>lt;sup>27</sup>The domain of the distribution needs to be a compact set when solving the model computationally. Hence, finite upper bounds are imposed on  $\boldsymbol{A}$  and  $\boldsymbol{Z}$ , and their values are set such that there is no mass on those points under the stationary distribution.

 $A \times Z \times H \times \Theta \times O$ . Two additional auxiliary objects are defined. First, occupational choice  $o'(a, z, h, \theta, o) : A \times Z \times H \times \Theta \times O \to O$  is defined as a function that solves (1.12) if o = w, and solves (1.14) and (1.16) if  $o \in \{er, ei\}$ . Second, the state vector for each occupation o is defined as  $\Omega_o = (a, z, h, \theta, o)$ .

It is assumed that new entrants to the economy begin as workers, and randomly draw assets a from the asset distribution in the previous period. Then the distribution  $\mu$  follows the law of motion  $\mu' = \Phi(\mu)$ , where

$$\mu'(\tilde{a}, \tilde{z}, h, \theta, \tilde{o}) = (1 - \zeta) \cdot \left\{ \int \int_{\mathbb{I}\{\tilde{a} = a'(\Omega_w), \tilde{z} = z, \tilde{o} = o'(\Omega_w)\}} \mu(\Omega_w) dadz + e^{-\Delta^R(\Omega_{er})} \int \int_{\mathbb{I}\{\tilde{a} = a'(\Omega_{er}), \tilde{z} = ze^{\Delta^R(\Omega_{er}), \tilde{o} = o'(\Omega_{er})\}}} \mu(\Omega_{er}) dadz + (1 - e^{-\Delta^R(\Omega_{er})}) \int \int_{\mathbb{I}\{\tilde{a} = a'(\Omega_{er}), \tilde{z} = ze^{-\Delta^R(\Omega_{er}), \tilde{o} = o'(\Omega_{er})\}}} \mu(\Omega_{er}) dadz + u(\Omega_{ei}) \int \int_{\mathbb{I}\{\tilde{a} = a'(\Omega_{ei}), \tilde{z} = ze^{\Delta^I(\Omega_{ei}), \tilde{o} = o'(\Omega_{ei})\}}} \mu(\Omega_{ei}) dadz + (1 - u(\Omega_{ei})) \int \int_{\mathbb{I}\{\tilde{a} = a'(\Omega_{ei}), \tilde{z} = z, \tilde{o} = o'(\Omega_{ei})\}} \mu(\Omega_{ei}) dadz \right\} + \zeta \cdot \int_{\mathbb{I}(o = w)} d\mu_a(a) dF(h) dG(z|h)$$

where

$$\mu_a(a) = \int \int \sum_{\theta \in \{0,\bar{\theta}\}} \sum_o \mu(\Omega_o) dz dh$$

A stationary recursive competitive equilibrium is defined as follows.

Definition For a given interest rate r, a stationary recursive competitive equilibrium is a set of value functions  $\{V^W, V^{E,R}, V^{E,I}\}$ , policy functions  $\{a', c, k, n, \Delta^R, u, o\}$ , wage rate w, and distribution  $\mu^*$  such that

#### 1. Individuals optimize:

 $V^W, V^{E,R}, V^{E,I}$  satisfy (1.11), (1.12), (1.13), (1.14), (1.15), and (1.16). Associated policy functions are as follows.  $a': A \times Z \times H \times \Theta \times O \to A$  is the savings decision;  $c: A \times Z \times H \times \Theta \times O \to \mathbb{R}^{++}$  is consumption;  $k: A \times Z \times H \times \Theta \times \{er, ei\} \to \mathbb{R}^+$  is capital demand;  $n: A \times Z \times H \times \Theta \times \{er, ei\} \to \mathbb{R}^+$  is labor demand;  $\Delta^R: A \times Z \times H \times \Theta \times \{er, ei\} \to \mathbb{R}^+$  is the risky experimentation choice;  $u: A \times Z \times H \times \Theta \times \{er, ei\} \to \mathbb{R}^+$  is the incremental innovation choice; and  $o: A \times Z \times H \times \Theta \times O \to O$  is the occupational choice.

#### 2. The labor market clears:

$$\sum_{\theta \in \{0,\bar{\theta}\}} \int n(\Omega_{er}) \mu(\Omega_{er}) da dz dh + \sum_{\theta \in \{0,\bar{\theta}\}} \int n(\Omega_{ei}) \mu(\Omega_{ei}) da dz dh$$
$$= \sum_{\theta \in \{0,\bar{\theta}\}} \int \int \int h \mu(\Omega_w) da dz dh$$

#### 3. Distribution is time-invariant:

$$\mu^*(a, z, h, \theta, o) = \Phi(\mu^*(a, z, h, \theta, o))$$

Luttmer (2012) shows that in a similar environment in which firm productivity processes follow a standard Brownian motion, a stationary distribution  $\mu^*$  exists as long as the death rate  $\zeta$  is large enough and the dispersion of entrants' productivity distribution is not too large. The intuition is that even though each cohort of innovating entrepreneurs moves upwards in productivity space  $\mathbf{Z}$ , their mass is reduced by rate  $\zeta$  every period and eventually converges to zero. At the same time,  $\zeta$  mass of individuals enter the economy at the lower part of the productivity distribution, balancing out overall growth in z achieved by entrepreneurs in the previous period. Exogenous churning induced by  $\zeta$  and entailed firm entry and exit costs prevent assets from diverging.

## 1.5 Quantitative Analysis

#### 1.5.1 Calibration

I calibrate model parameters to match certain key features of the U.S. non-farm private sector between 1999 and 2014. This period is chosen so that I can exploit the regression results obtained in Section 1.3 to discipline the model. A subset of parameters are fixed at values commonly used in the macroeconomics literature. The remaining parameters are chosen to minimize the distance between a set of equilibrium moments obtained from model simulation and their data counterparts. The model parameters are summarized in Table 1.9 and Table 1.10.

Externally Calibrated Parameters The model period is equivalent to one year. I set the time discount factor  $\beta$  to 0.968. The relative risk-aversion coefficient  $\sigma$  is set to 2, which is standard in the literature. I set the decreasing returns to scale parameter  $\nu$  to 0.85, as in Atkeson and Kehoe (2007) and Midrigan and Xu (2014). I set the capital depreciation rate  $\delta$  to 0.065 and the interest rate to 0.03. The R&D cost elasticity parameter  $\psi$  is set at 2, which is standard in the endogenous growth and innovation literature (e.g., see Acemoglu, Akcigit, Bloom, and Kerr, 2013 and Akcigit and Kerr, forthcoming). Labor efficiency depreciation rate of entrepreneurs,  $\delta_h$ , is set at 0.03. This value is taken from existing empirical studies which find that returns to entrepreneurship experience are on average lower than the returns to experience as paid workers.<sup>28</sup> 3% is within the range of estimates reported in the literature. I set the labor efficiency distribution dispersion parameter  $\eta_h$  to 1.41, to match the 90th to 10th percentile ratio of the weekly earnings distribution reported by the BLS. The ratio is calculated as the average value between 1999 and 2014.<sup>29</sup>

Table 1.9: Externally Calibrated Parameters

Parameter		Value	Source
Discount factor	β	0.968	Standard
Risk-aversion coefficient	$\sigma$	2	Standard
Capital production share	$\alpha$	0.330	Standard
Returns to scale parameter	$\nu$	0.850	Standard
Capital depreciation rate	$\delta$	0.065	Standard
Interest rate	r	0.030	Standard
R&D cost elasticity	$\psi$	2	Akcigit and Kerr (2017)
Human capital depreciation rate	$\delta_h$	0.030	Literature
Labor efficiency dispersion	$\eta_h$	1.410	p90/p10 week. earn. 1999-2014 (BLS)

<sup>&</sup>lt;sup>28</sup>For example, see Bruce and Schuetze (2004), Kaiser and Malchow-Møller (2011), and Baptista, Lima, and Preto, 2012.

<sup>&</sup>lt;sup>29</sup>The BLS computes the weekly earnings distribution percentiles from the Current Population Survey (CPS) sample that comprises of wages and salary workers who are 25 years or older.

Internally Calibrated Parameters The exogenous death rate  $\zeta$  is set to 0.05 to reproduce the exit rate of mature firms (age 6 or higher) observed in the data, where the data moment is calculated as the average value between 1999 and 2014 from the Census Bureau's Business Dynamics Statistics (BDS). The implied effective discount rate  $\beta(1-\zeta)$  is 0.92, which is consistent with the choice of Buera, Kaboski, and Shin (2011) and Midrigan and Xu (2014). The financial frictions parameter,  $\lambda$ , is set to 4.5 to match the average value of  $\frac{k-m}{k}$  of all firms in the model to the average ratio of liabilities to nonfinancial assets for the U.S. nonfinancial business sector between 1999 and 2014.<sup>30</sup>

Hurst and Pugsley (2011) document that about 50% of startup entrepreneurs in the Panel Study of Entrepreneurial Dynamics (PSED) report nonpecuniary motives as one of the primary reasons for starting their businesses, and the majority of those entrepreneurs' firms remained small throughout the sample period of their study. In the model, higher values of  $\theta$  induces individuals to enter entrepreneurship with smaller sizes and to take smaller risks as shown in Prediction 4 in Section 1.2. Thus, I jointly calibrate  $\theta$  and  $p_{\theta}$  to replicate that startup entrepreneurs with  $\theta = \bar{\theta}$  comprise 50% of all startup entrepreneurs in the model, and to target the share of small firms (less than 20 employees) in the economy.

The initial distribution of z is governed by three parameters:  $\underline{z}$ ,  $\sigma_z$ , and  $\rho$ . I

<sup>&</sup>lt;sup>30</sup>I follow Buera and Nicolini (2017) and compute the statistics from the U.S. flow of funds. I measure liabilities as the sum of total liabilities of noncorporate (FL114190005.Q) and corporate (FL104190005.Q) firms in the nonfinancial sector minus the U.S. real estate owned by foreigners (FL115114005.Q) and the foreign direct investment in the U.S. (FL103192005.Q). Similarly, I measure nonfinancial assets as the sum of nonfinancial assets of noncorporate (FL112010005.Q) and corporate (FL102010005.Q) firms in the nonfinancial sector minus the U.S. real estate owned by foreigners (FL115114005.Q) and the foreign direct investment in the U.S. (FL103192005.Q).

calibrate  $\underline{z}$  and  $\sigma_z$  to match the (employment-weighted) firm entry rate and the ratio between the average employment of entrants relative to the average employment of incumbents, respectively. To discipline  $\rho$ , I use the micro-level relationship between startup employment and log prior earnings reported in Table 1.3. I normalize startup employment by the average employment over all firms calculated from the BDS, which is 22.8, and obtain a normalized coefficient of 0.105. I simulate the model and create a cohort of 100,000 startup firms, and run an identical regression with the simulated data to calibrate  $\rho$ , in which employment levels of startups in the simulated data are normalized by average employment over all firms in the model.

To calibrate the elasticity of the innovation success probability with respect to risk choice,  $\gamma$ , I run a regression with the simulated data that is counterpart to the regression of exit on log prior earnings and firm characteristics reported in Table 1.4. Because the model describes a single good economy with no price heterogeneity and adjustment frictions, the simulated data do not show variation in revenue labor productivity, as in the empirical data. Therefore, I use z in the simulated data regression. I target the coefficient on log prior earnings of 0.019 reported in columns (3) and (4) of Table 1.4. I set the experimentation cost parameter F, another factor that strongly governs the risk-taking incentives of young firms, to match the average employment growth rate of young firms. Because the exit rates of old firms are determined by the exogenous death rate  $\zeta$ , the fixed firm entry and exit cost  $\phi$  mostly governs the exit rate of young firms. Hence, I calibrate  $\phi$  to match the average exit rate of young firms. In the model, most incremental innovation is conducted by old firms whose z's have moved sufficiently far away from their exit margin.

Therefore, I target the employment growth of old firms (age 10+) to calibrate the incremental innovation step size  $\Delta^{I}$ . I choose the incremental innovation research cost scale parameter,  $\chi$ , to match the R&D intensity of innovating firms documented by Akcigit and Kerr (forthcoming).

Table 1.10: Internally Calibrated Parameters

Parameter		Value	Value Target	Source	Model Data	Data
Financial friction	~	4.5	Liability to nonfinancial assets	U.S. Flow of Funds	0.707	0.695
Death probability	Ş	0.05	Age 6+ exit rate	BDS 1999-2014	0.053	0.059
z distribution location	54	1.0	Firm entry rate	BDS 1999-2014	0.021	0.025
$z$ distribution dispersion $\sigma_z$	$\mathcal{I}_z$	9.0	Emp.ratio entrants to incumbents	Startup Data	0.331	0.270
z and $h$ dependence	θ	0.09	Reg. age 0 emp. vs. $\ln(PrEarn)$	Startup Data	0.097	0.102
Risk prob. elasticity	~	1.2	Reg. exit. vs. $\ln(PrEarn)$	Startup Data	0.029	0.019
Risky experiment cost	F	0.16	Age 1-5 avg emp. growth	Startup Data	0.089	0.065
Firm entry / exit cost	φ	0.1	Age 1-5 exit rate	Startup Data	0.104	0.134
Increm. innov. step size $\Delta^I$	7	0.3	Age $10+$ avg emp. growth	Acemoglu et al (2013)	0.034	0.015
Increm. innov. cost scale $\chi$	$\times$	3.5	R&D - Sales ratio	Akcigit & Kerr (2017)	0.010	0.042
Preference for entrep.	$\bar{ heta}$	8.0	Small (emp $\leq 20$ ) firm share	BDS 1999-2014	0.700	0.890
$\operatorname{Prob}(\theta = \overline{\theta})$ 1	$\theta a$	0.11	Frac. nonpecuniary entrep.	Hurst & Pugsley (2011)	0.498	0.500

## 1.5.2 Calibrated Model Properties

In this section, I document the key properties of the calibrated model. Figure 1.3 plots the average employment of each firm age group computed from the BDS and those calculated using the data obtained from model simulation. The BDS statistics are calculated as the average values between 1999 and 2014. Since employment in the model is expressed in the effective unit of labor, employment levels in the BDS and the model are not directly comparable. Therefore, I normalize the entrants' average employment in both data series to one and compare the slope over the life cycle. As shown in the figure, the model does a reasonable job in tracking the average size by age observed in the empirical data.

4.5
4
3.5
2
1.5
1
0.5
0
1
2
3
4
5
6 to 10 11 to 15 16 to 2021 to 25

Data Model

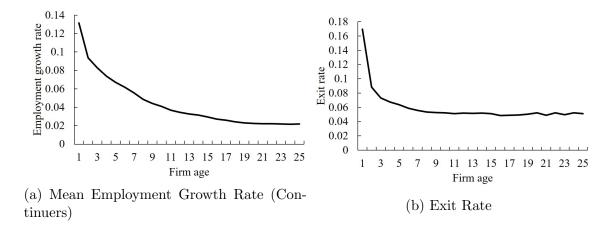
Figure 1.3: Average Employment by Firm Age

**Notes:** The data corresponds to the average employment by firm age from the Business Dynamics Statistics. The data values are computed as the average values between 1999 and 2014. Model statistics are calculated from a simulated data that contains a cohort of 100,000 startup firms.

Figure 1.4 shows the mean employment growth rate of the continuing firms and exit rates over firm age in the model-simulated data. Both series exhibit a convex

decreasing shape: younger firms exit at a higher rate, but conditioning on survival, they grow faster. This up-or-out growth dynamics of young firms implied by the model is consistent with the empirical findings in the literature (e.g., see Dunne, Roberts, and Samuelson, 1989; Evans, 1987; Haltiwanger, Jarmin, and Miranda, 2013).

Figure 1.4: Life-Cycle Growth and Survival Dynamics in the Model



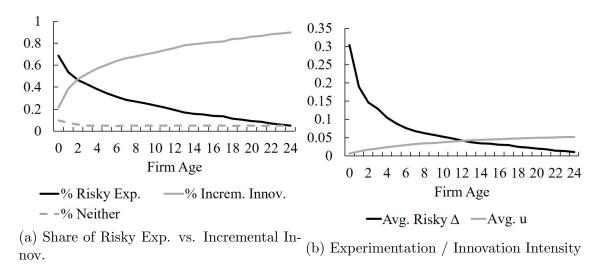
In this model, the decline in exit and employment growth rates with respect to firm age is driven by the higher intensity of risk-taking behavior by young firms. Young firms tend to take larger risks because of the following reasons: First, entrepreneurs start their firms with lower levels of productivity compared to those of the incumbents, and secondly, labor market clearing condition determines the wage such that startup entrepreneurs are close to their occupation switching margin. Therefore, returning to paid employment is a viable exit option for startup entrepreneurs in the case of failure from risky experimentation, which incentives them to take larger risks as shown in Section 1.2. As firms get older, risk-taking winners achieve an increase in their business productivity and risk-taking losers

either shrink or exit. Therefore, the productivity levels of continuing firms gradually move away from the exit threshold and the entrepreneurs conduct less risky experimentation as their firms age.

In Figure 1.5, I plot the experimentation and innovation patterns of continuing firms over their life cycle. Figure 1.5a shows the fraction of entrepreneurs who engage in risky experimentation, or incremental innovation, or neither. Figure 1.5b shows the average choice of  $\Delta^R$  conditioning on  $\Delta^R > 0$ , and the average incremental innovation success probability u conditioning on u > 0. The fraction of entrepreneurs conducting risky experimentation and their risk-taking intensity declines as their firms age. Simultaneously, entrepreneurs gradually switch over to incremental innovation. Also in the early phase of the firm life cycle, a significant fraction of entrepreneurs do not engage in any innovation activities. For instance, at age zero, when the risk-taking incentives are the greatest, about 10% of entrepreneurs do not conduct any experimentation or innovation and 20% of entrepreneurs exert only negligible effort in incremental innovation (average u of 0.006). Therefore, about in total 30% of firms show little or no growth in productivity at age one.

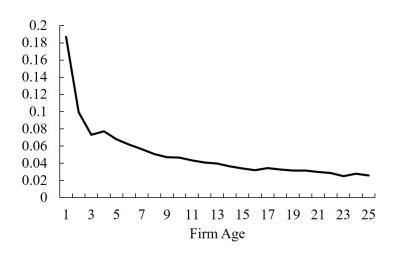
A higher intensity of risky experimentation results in a more rapid pace of selection and reallocation, which in turn drives up the average productivity of the continuing firms. Since the intensity of risky experimentation is higher for younger firms, growth in average productivity declines in firm age as illustrated in Figure 1.6. This model implication is consistent with the recent empirical findings of Alon, Berger, Dent, and Pugsley (2017) where they show in the U.S nonfarm business sector, the relationship between firm age and productivity growth is downward sloping

Figure 1.5: Experimentation and Innovation of Continuing Firms



and convex, and that most of the productivity growth is concentrated among firms less than five years old.

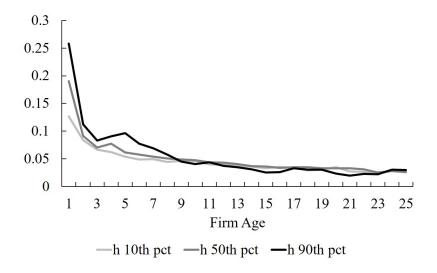
Figure 1.6: Growth of Average Productivity by Firm Age



In addition, since entrepreneurs with better outside options tend to take larger risks as shown in Section 1.2 and 1.3, average productivity of the firms operated by those entrepreneurs will grow faster. In Figure 1.7, I plot growth in average

productivity for firms operated by entrepreneurs with h in the 10th, 50th, and 90th percentile in the h distribution. The figure shows faster growth in average productivity for firms operated by a higher h.

Figure 1.7: Growth of Average Productivity by Firm Age with Different Outside Options



## 1.5.3 Counterfactual Exercises

# 1.5.3.1 Removing the Outside Option

To study the quantitative importance of outside options and the associated risk-taking behavior of young firms in the aggregate economy, I study a counterfactual situation in which entrepreneurs cannot return to paid employment. Though this is an extreme experiment, it provides a useful insight on how the existence of outside options affects the composition of startup firms and their life-cycle dynamics. It also provides an estimate of the upper bound of the output and productivity

losses an economy can suffer from overall deterioration of the outside options of startup entrepreneurs.

Table 1.11 shows a comparison of key statistics between the benchmark economy and the counterfactual economy. By construction, the firm exit rate falls to the exogenous death rate in the counterfactual case. Interestingly, the average productivity and size of entrants increase in the counterfactual economy. This is because if individuals know that they can never go back to paid employment, they will enter entrepreneurship only if their initial business productivity endowments are high enough. This positive selection effect, together with the steady-state force which equates the entry rate to the exit rate, induces the firm entry rate to fall. This result also indicates that an overall decline in outside options generates fewer but initially better startup firms.

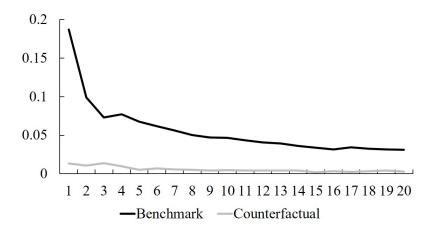
Table 1.11: Overall Effect of Removing the Outside Option

	Benchmark	Counterfactual	% Changes
Firm entry rate	0.065	0.050	-23%
Firm exit rate (age 1-5)	0.104	0.050	-51.9%
Firm exit rate (age $6+$ )	0.055	0.050	-9.1%
Average entrant size (h)	5.775	6.576	13.9%
Average entrant z	4.403	5.314	20.7%
Aggregate output	2.797	2.547	-8.9%
Aggregate output per worker	1.438	1.375	-4.4%

However, aggregate output and labor productivity fall significantly in the counterfactual economy, by 8.9% and 4.4%, respectively. This is driven by differences in the growth rates of young firms along their life cycle. Because entrepreneurs in the counterfactual economy do not have an exit option to exercise in the case of business failure, they do not take risks early in their life cycle and thus show little

or no growth as a group. On the other hand, although startups in the benchmark economy begin with lower productivity, they take larger risks and thus grow much faster as they age. Figure 1.8 illustrates this point by plotting the growth rates of the average productivity by firm age group. Therefore, this counterfactual experiment reveals that deterioration (or, for the same reason, improvement) of the outside options of startup entrepreneurs can have a sizable impact on aggregate output and productivity.

Figure 1.8: Growth of Average Productivity by Firm Age: Benchmark vs. Counterfactual



**Notes:** This figure shows the growth rate of average productivity by firm age in the benchmark economy and counterfactual economy where the outside options of the entrepreneurs are removed. Statistics are calculated from panel data with 100,000 startup firms obtained from the model simulation.

## 1.6 Conclusion

In this paper I show that the outside options of startup entrepreneurs, which I define as the level of labor income they expect to earn in the case of business failure, are a key predictor of the early growth trajectories of young firms. Better outside

options serve as an effective channel of insurance against business failure, which enables entrepreneurs to take larger business risks. Larger risk-taking behavior translates into a more dispersed, up-or-out type of firm dynamics. I test these predictions using a large founder-firm matched administrative data set and find that the model implications are empirically supported.

I also show that large changes in the outside options of startup entrepreneurs can potentially have a large impact on aggregate output and productivity. An improvement in outside options induces smaller and less productive firms to enter, but incentivizes them to engage in riskier experimentation and exhibit faster average productivity growth along their life cycle. Therefore, the post-failure options of entrepreneurs are an important factor that governs not only young firm growth and survival, but also aggregate output and productivity growth.

The quantitative framework established in this paper can be extended to study several critical questions posed in macroeconomics. First, the model can be used to study whether the rise in labor income inequality in the U.S. during the last three decades have had an impact on the decline in high-growth entrepreneurship and business dynamism (Decker, Haltiwanger, Jarmin, and Miranda, 2016). While individuals with higher labor earnings are more likely to create high-growth young firms, it is possible that the rapid increase in their labor earnings may have made them less likely to enter entrepreneurship in the first place. Second, one can study the business cycle implications of this mechanism by introducing aggregate uncertainty and unemployment shocks in the model. When the aggregate economy is in a downturn and the unemployment rate is high, startup entrepreneurs would expect

to experience difficulties finding a job if they cease their business operation. This would incentivize them to take less risks, further reducing young firm growth and thus resulting in lower levels of job creation in the economy. Therefore, risk-taking by startup entrepreneurs may work as a propagation mechanism of aggregate shocks.

Chapter 2: Lobbying, Procurement Allocation, and the Employment

Effect of Fiscal Stimulus (coauthored with Veronika Pen
ciakova and Felipe Saffie)

#### 2.1 Introduction

Fiscal stimulus packages aim to stabilize employment and output during crises, and their effectiveness remains an open question. The enactment of the \$832 billion American Recovery and Reinvestment Act (ARRA) in the midst of the Great Recession in February 2009 renewed interest in the impact of stimulus spending. One of the primary goals of ARRA was to save or create up to 3.5 million jobs, with over 90 percent of those jobs being in the private sector. Recent empirical macroeconomics literature has primarily focused on the employment effect of aid to state governments (Chodorow-Reich, Feiveson, Liscow, and Woolston, 2012) and government purchases (Dube, Kaplan, and Zipperer, 2014; Dupor and Mehkari, 2016; Feyrer, 2011; Wilson, 2012). We take an additional step forward and ask how stimulus spending was allocated, and whether its impact on employment depends on who it was allocated to.

We focus on the allocation and impact of one important component of ARRA

spending: government purchases through federal procurement. Between 2009 and 2014, ARRA-supported federal procurement contracts amounted to \$130 billion. Deteriorating economic conditions, combined with this influx of money, may have generated incentives for firms to try to influence the allocation of these contracts. In this paper, we emphasize political influence through lobbying. We know that lobbying firms are active participants in the Federal procurement market. Each year, they account for 1.5% of the 165,000 procurement contractors and 55% of total procurement spending. We therefore ask whether corporate lobbying influenced the allocation of contracts during the crisis; and whether contracts subject to lobby have a differential impact on employment outcomes.

To answer the first question we need to identify the causal relationship between lobbying and procurement outcomes. There are several threats to identification. First, in the data it is challenging to link lobbying activity on particular issues with expenditure on particular procurement contracts. Second, lobbying is an endogenous choice made by firms that may be driven by observable (e.g., firm size, industry) and unobservable characteristics (e.g. political connections). Fortunately, ARRA provides a suitable laboratory.

In particular, the richness of the procurement and lobbying data allows us to identify both lobbying activity associated with ARRA and the allocation of individual procurement contracts supported by ARRA at the firm level. Additionally, the swift introduction and passage of the stimulus bill assuage concerns regarding lobbying behavior in previous years being targeted toward ARRA contracts; these contracts after all did not exist prior to the passage of the bill. In fact, the first

version of ARRA was introduced in early January 2009, and was signed into law less than two months later in February.

To resolve the challenges associated with selection into lobbying on ARRA, we first prune the sample by matching on a number of observable characteristics. These include firm size, industry, prior lobbying status, and both the number and average size of past procurement contracts. After matching, we obtain 1,061 firms that lobbied on ARRA (treated) matched with 498 similar firms that did not lobby (control). We validate our pruning strategy by ensuring that our resulting treatment and control groups are similar in all observable dimensions and that they have similar pre-ARRA procurement outcomes. We then assess the impact of lobbying on the allocation of ARRA-supported procurement contracts and find that firms that lobby on ARRA are more likely to win these contracts and win 5.3% more and 50% larger ARRA contracts.

This matching approach may not fully capture all the unobserved heterogeneity. For example, even after matching, firms that lobby on ARRA could be more politically connected to the government. In such cases, firms that lobby on ARRA should win more procurement contracts in general, and we might wrongfully attribute this effect to lobby. We validate our identification strategy by showing that ARRA lobbying only has an effect on ARRA procurement outcomes and does not have any significant impact on the corresponding non-ARRA outcomes.

A critical question remains. Did contracts subject to lobbying on ARRA have the same effect on local employment as other procurement contracts? To answer this question we build on the rich literature assessing the employment effect of stimulus spending (Chodorow-Reich, Feiveson, Liscow, and Woolston, 2012; Dube, Kaplan, and Zipperer, 2014; Dupor and Mehkari, 2016; Feyrer, 2011; Wilson, 2012). We establish causal relationship between ARRA spending and employment growth outcomes in Metropolitan Statistical Area (MSA) by constructing a set of Bartik instruments.

We take advantage of pre-ARRA (2007) procurement information and aggregate to the MSA-sector level to determine how reliant procurement in each sector is on each MSA. We assume this reliance is relatively sticky and use it to allocate national ARRA spending in a particular sector across MSAs. By aggregating over all the sectors within a MSA, we get an instrument for total ARRA spending. To instrument for the share of procurement spending through lobbying firms, we take advantage of persistence in corporate lobbying (Kerr, Lincoln, and Mishra, 2014a) and variation in the intensity of procurement through lobbying firms across sectors. Over 90% of firms that lobbied on ARRA in 2009 had lobbied on other issues in the previous three years. And while some sectors channel over 80% of total procurement through lobbying firms, others channel as little as 1\%. We first calculate the pre-ARRA (2007) share of procurement spending in each sector channeled through firms that lobbied on budget and appropriation. This share is then combined with our MSA-sector reliance measure and national sector-level ARRA spending. The resulting MSA-sector estimate of ARRA spending through lobbying firms is aggregated to the MSA level to obtain our second instrument.

After constructing our instruments, we first estimate that ARRA procurement spending yields 11.5 jobs per \$1 million spent. This estimate are broadly consistent

with the fiscal multiplier reported in Dube, Kaplan, and Zipperer (2014) who reports 7.5 job-years created per \$1 million spent, or Wilson (2012) who reports eight jobs. We then disaggregate total ARRA procurement per capita into the amount channeled through firms that lobbied on ARRA and those that did not. We find a striking result. The effect of ARRA procurement spending on employment is entirely driven by non-lobbying firms. Stimulus money channeled through these firms yields 16 jobs per \$1 million spent, while money channeled through lobbying firms has no effect on employment. Although this estimate does not take into account general equilibrium effects, it does suggest that the impact of stimulus spending on employment is attenuated by the allocation of contracts to lobbying firms. In short, it is not only the amount, but also the allocation of stimulus spending across firms that matter for employment outcomes.

The paper is organized as follows. Section 2.2 reviews the related literature. Section 2.3 describes the goals and characteristics of the American Recovery and Reinvestment Act. Section 2.4 describes the construction of the data. Section 2.5 discusses the impact of lobbying on procurement allocation. Section 2.6 documents the differential effects of ARRA on employment. Finally, section 2.8 concludes.

#### 2.2 Literature Review

This paper belongs at the intersection of three strands of literature. Our focus on whether and how lobbying affects the allocation of procurement spending connects us to the literature investigating various mechanisms through which procurement spending is allocated to firms. The empirical literature on public procurement examines the factors that shape competitiveness and contractual terms (Bajari, McMillan, and Tadelis, 2008; Bajari and Tadelis, 2001; Warren, 2014), and how different designs of procurement processes affect their efficient allocation outcomes (see Bhattacharya and Sweeting (2015) for a review). Further, Liebman and Mahoney (2013) show that wasteful year-end fiscal spending leads to inefficient procurement allocation in general. In contrast to these studies, we examine lobbying as a determinant of procurement allocation.

Our emphasis on the effect of corporate lobbying on the allocation of procurement contracts during the stimulus closely connects us to the broader literature on the implications of corporate lobbying. This literature has primarily focused on preferential tax treatment (Arayavechkit, Saffie, and Shin, 2014; Meade and Li, 2015; Richter, Samphantharak, and Timmons, 2009) and trade policy (Bombardini, 2008; Bombardini and Trebbi, 2012; Gawande and Bandyopadhyay, 2000). Lobbying has also been found to increase the likelihood of receiving government relief (Adelino and Dinc, 2014; Blau, Brough, and Thomas, 2013; Duchin and Sosyura, 2012) and to generate high returns when policies are enacted (Kang, 2015). With the exceptions of Brogaard, Denes, and Duchin (2016) and Adelino and Dinc (2014), there are few examinations of how corporate lobbying affects the allocation of procurement contracts.

Whereas Brogaard, Denes, and Duchin (2016) identifies the positive effect of corporate political connections on the allocation of procurement contracts by exploiting campaign contributions (PAC) in close elections, we use corporate lobbying.

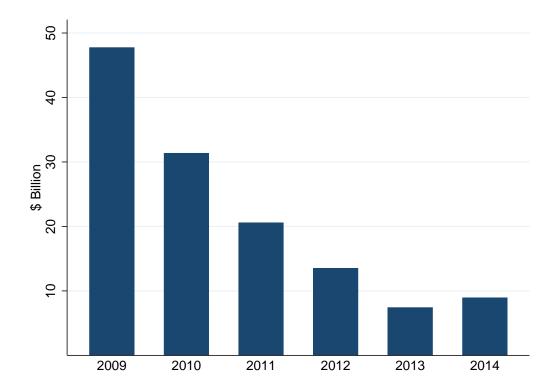
The use of lobbying affords us two advantages. Corporate lobbying expenditure is seven times larger than corporate campaign contributions, and we can directly link lobby expenditure to particular bills. Adelino and Dinc (2014) document a positive correlation between lobbying and the receipt of ARRA stimulus funds. By linking of ARRA-related lobbying with ARRA-related procurement and correcting for selection bias, we take the analysis one step further and establish the causal link between lobbying and the allocation of procurement contracts.

Finally, by assessing the importance of procurement allocation for real outcomes, we contribute to the empirical literature evaluating the effectiveness of stimulus spending. The recent literature in this area has focused primarily on ARRA. Some consider the bill's provisions regarding aid to state governments (Chodorow-Reich, Feiveson, Liscow, and Woolston, 2012; Wilson, 2012), while others emphasize the provisions for low-income households, infrastructure spending (Feyrer, 2011), and total government purchases (Dube, Kaplan, and Zipperer, 2014; Dupor and Mehkari, 2016; Wilson, 2012). This strand of literature is principally concerned with estimating the local employment multiplier. Dube, Kaplan, and Zipperer (2014) estimate an annualized employment multiplier of 7.5 job-years per \$1 million, which is close to the eight jobs estimated by Wilson (2012). Meanwhile, Dupor and Mehkari (2016) estimates that ARRA increases local employment by 9.53 persons at the commuting zone level. We use insights from these papers regarding the regression framework and appropriate controls to ensure that our approach and results are comparable to previous work. We contribute to this literature by focusing on an understudied, yet important, category of government purchases during stimulus - Federal procurement. More importantly, we emphasize the differential effect of spending through lobbying and non-lobbying firms on local employment outcomes.

### 2.3 The American Recovery and Reinvestment Act

To assess the impact of political influence on the allocation and employment effect of procurement contracts during periods of fiscal stimulus, we consider the passage of the American Recovery and Reinvestment Act. ARRA was passed in February 2009 in response to the Great Recession with a stated goal of stabilizing the economy and saving jobs through temporary relief programs, federal tax incentives, and government purchases. Of the estimated \$831 billion to be spent beginning in 2009, 31% was allocated to loans, grants, and procurement for infrastructure, energy, communications and scientific research. In total between 2009 and 2014, approximately \$130 billion was allocated towards federal procurement. The majority of this contracting occurred between 2009 and 2010. As figure 2.1 shows, by the end of 2010, 61% (\$79 billion) of contract dollars had been spent.

ARRA was first introduced in the Senate on January 6, 2009 and in the House on January 26. It was signed into law by President Obama on February 17. The swift introduction and passage of ARRA reduce the likelihood that firms changed their behavior in anticipation of receiving ARRA-supported contracts, which would bias our empirical results. ARRA's \$831 billion price tag made it the largest single stimulus bill in U.S. history and efforts were made to monitor spending. In particular, legislation required that all awards funded by ARRA be reported on quarterly



**Notes:** This figure shows the total spending on ARRA supported procurement between 2009 and 2014. It The totals include initial obligations, as well as follow-up contract modifications.

Figure 2.1: Arra Contracts: Spending by Year (2009-2014)

and that these reports be made public. Our empirical strategy takes advantage of this publicly available information to identify procurement contracts associated with ARRA. We then integrate publicly available information on lobbying to identify procurement contractors that lobbied on ARRA. Our ability to both identify ARRA procurement and ARRA lobbying allows us to more tightly isolate how political influence, exerted through corporate lobbying, affects the allocation and local employment impact of federal procurement during the stimulus.

### 2.4 Data Description

#### 2.4.1 Firm-level analysis

Our empirical analysis relies on several sources. Federal procurement data is obtained from USAspending.gov, a website mandated by the Federal Funding Accountability and Transparency Act of 2006. This site hosts data on the universe of federal procurement contracts awarded since 2000, with size above \$3,000 (\$25,000 prior to 2005). The data contain detailed contract-level information, including contract size, terms, awarding agency, location of performance, and product/service type. Additionally, information is available on the recipient, including business name, location and employment.

Among all the contracts awarded, we identify the ARRA-supported contracts using the Recovery Report data from the Federal Procurement Data System (FPDS). Section 1512 of the ARRA requires that recipients of ARRA resources report certain information such as the amount of recovery funds received and a list of projects for which funds will be used. Further, government agencies are required to review the Recovery Report posted on the FPDS website every day to ensure that all entries are accurate. We match the USAspending.gov data and Recovery Report to identify ARRA supported contracts among all procurement contracts. Linking ARRA contracts to the universe of federal procurement contracts is central to our empirical strategy. By doing so, in our firm-level analysis we can control for past procurement experience, and in our Metropolitan Statistical Area (MSA)analysis

we can use pre-ARRA procurement in constructing our Bartik instruments.

The federal procurement data are used to construct several outcome variables of interest. For our firm-level analysis we construct a dummy equal to one if the contract is associated with ARRA (DARRA), total number of ARRA contracts (Narra), total first-year value of ARRA contracts (Varra), total number of non-ARRA contracts (NnonARRA), and total first-year value of non-ARRA contracts (VnonARRA). When validating our matching strategy we also use the fraction of contracts awarded competitively (COMP), where competitive contracts are defined as those classified as being awarded under full and open competition. We also construct key control variables: log average first-year value of new contracts in the previous three years (MP3) and log total number of new contracts won in the previous three years (NP3), employment (EMP), and 2-digit industry (NAICS). Because our employment data is derived from reports filed by contracting firms, which is not as reliable as employment from administrative records, rather than use a continuous measure of employment, we place firms into four bins. The first bin contains firms with less than 50 employees, the second with 50 to 249 employees, the third with 250 to 999 employees and the fourth with 1,000 or more employees.

Lobbying data are obtained from the Center for Responsive Politics (CRP). The Lobbying Disclosure Act of 1995 requires the disclosure of lobbying activity to the Clerk of the U.S. House of Representatives and Secretary of the U.S. Senate when expenditure exceeds \$3,000 during a quarter. In addition to the total amount of expenditure, disclosures also report which issue areas and bills were targeted in lobbying efforts. The data also include the name of organizations or firms on

behalf of which lobbying is done. We use this information and a probabilistic name matching algorithm to link lobby data to federal procurement data.

In particular, we first standardize names in the procurement and lobbying data to eliminate punctuation, legal form information, and make adjustments for common acronyms. The names in each data set are further processed to generate match-codes. The procurement and lobby data are then matched based on these codes. More often than not, each entity in the procurement data will at first be matched to multiple entities in the lobby data. We use a Jaro-Winkler distance score to evaluate matches and for each firm keep the match with the highest score. Among the remaining matches, those with a low score are also dropped. We end up finding approximately 3% of contractors in the lobby data, and conversely, contractors account for nearly 30% of lobbying entities during the period 2008 through 2015.

In our empirical analysis we use two variables derived from the lobbying data. The first is a time-varying dummy equal to one if a firm lobbied on any issue in the previous three years (LP3) and the second is a time-invariant dummy equal to one if a firm lobbied on ARRA (LARRA). We identify a firm as lobbying on ARRA if during the 111th Congress it lobbied on any of the House or Senate versions of ARRA, or related bills (H.R.1 and S.1). Lobbying on these bills is identified in the cleaned bill-level data available from CRP.

Our firm-level data set contains a little over 1 million procurement contracts awarded between 2008 and 2015 to 306,000 contractors. Around 42,000 of these are ARRA-supported contracts awarded to over 10,000 contractors. A total of 5,000 firms in the data lobby between 2008 and 2015, while only 850 (17%) of these lobby

specifically on ARRA. Firms that lobby on ARRA represent only 1.5% of the total number of firms awarded ARRA contract, yet they account for around 34% of ARRA contracts and around 30% of ARRA contract spending between 2009 and 2015.

### 2.4.2 MSA-level analysis

For our MSA-level analysis, we use the procurement data to construct aggregate measures of procurement. Since we are interested in identifying the impact of spending on local employment outcomes, we focus on the location of performance, rather than the location of firms. Further, we measure spending as the sum of all spending obligations, which include initial contract value along with all subsequent contract modifications.<sup>1</sup> We calculate total ARRA spending by MSA and three-digit NAICS (ARRA), as well as total ARRA spending channeled through firms that lobbied on ARRA (ALOB). Our Bartik instruments are built using pre-ARRA procurement information. At the national level, we measure the share of 2007 procurement spending in each sector going to each MSA (ProcShare), as well as the total share of 2007 procurement spending in each sector channeled through firms that lobbied on budget and appropriations in 2006 or 2007 (BLOB).

We incorporate additional data sources to generate control variables previously emphasized in the literature. Our measure of 2009 working age population (*POP*) is obtained from the U.S. Census Bureau. MSA-level quarterly employment information is obtained from the Quarterly Census of Employment and Wages (QCEW)

<sup>&</sup>lt;sup>1</sup>Contract modifications involve additional work, changes in costs, termination, etc. We do not distinguish between the type of modification when calculating spending.

hosted by the Bureau of Labor Statistics (BLS). Using this information, we calculate the 2009Q1 employment to population ratio ( $\frac{EMP}{POP}$ ). These data are also used in the construction of a Bartik measure of predicted change in employment between 2009Q1 and 2011Q1, as emphasized by Dube, Kaplan, and Zipperer (2014). We also control for the change in housing price index between 2003 and 2007 (HPI), using data from the Federal Housing and Finance Agency (FHFA); the share of less-educated young men in 2008 (SLME) from the American Community Survey (ACS); 2008 unemployment rate (UNEMP) from the BLS; and a dummy variable identifying whether a county in the MSA was represented by a congressperson on a committee that ARRA legislation went through.

### 2.5 Lobbying and ARRA Procurement

### 2.5.1 Descriptive Statistics

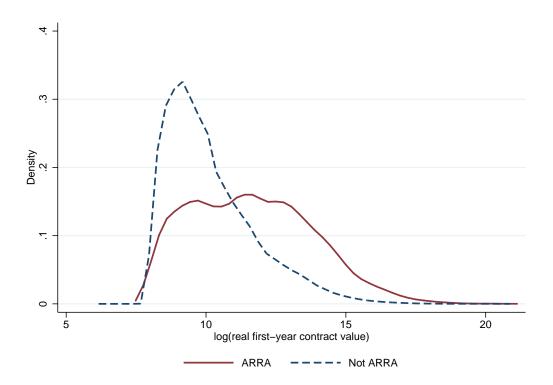
Before formally identifying the causal impact of ARRA lobbying on the allocation of ARRA contracts, let us explore what the raw data suggest about this relationship. The lobbying data show that the American Recovery and Reconstruction Act attracted significant attention from lobbying firms. In 2009, approximately \$2.4 billion were spent on lobbying and ARRA-related lobby alone accounted for around 7% of the total expenditure.<sup>2</sup> A significant fraction of the entities lobbying to shape ARRA were firms active in the procurement market. Among the 2,100 entities

<sup>&</sup>lt;sup>2</sup>Lobbying disclosure requirements do not require firms to report expenditure separately for each bill. As is standard in this literature, we divide expenditure equally across all bills listed in each disclosure. As such, we anticipate that the ARRA lobby expenditure reported here is a lower bound.

(e.g. foreign governments, associations of firms, associations of consumers, public entities, and companies) 850 (40%) are procurement contractors. These contractors account for two-thirds of all ARRA lobbying expenditure. Nevertheless, lobbying on ARRA-related bills does not guarantee that firms win ARRA contracts. Among the contractors that lobbied on ARRA, approximately 18% were awarded at least one ARRA contract between 2009 and 2014.

Yet, the returns to winning contracts appear quite high. Figure 2.2 compares the size distribution of the average ARRA and non-ARRA contract awarded to firms, where size is measured as the first year dollar value of a contract. To be clear, if a firm has two ARRA contracts and three non-ARRA contracts, the average first-year value of the first two contracts is used to build the distribution of ARRA contracts and the average of the other three is used to build the distribution of non-ARRA contracts. Every firm in the procurement database with active contracts is used to generate this figure. The average ARRA contract awarded to firms is on average larger than the average non-ARRA contract awarded to firms. In fact, the average size of an ARRA contract is nearly \$1.6 million compared to \$250,000 for non-ARRA contracts.

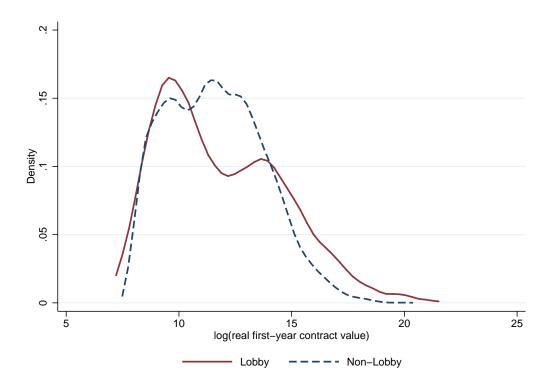
Figure 2.3 divides the aforementioned size distribution of ARRA contracts between contractors that lobbied on ARRA and those that did not. Figure 2.3 suggests that large ARRA contracts are more likely to be awarded to firms that lobbied on ARRA. The average ARRA contract size to non-lobbying firms is \$1.3 million compared to \$8.6 million to lobbying firms. There is a similar fraction of lobbying and non-lobbying firms receiving small contracts. But, non-lobbying firms



**Notes:** This figure shows distribution of log mean first-year contract value for ARRA and non-ARRA contracts at the firm level between 2009 and 2015. The figure excludes observations with zero contract value.

Figure 2.2: Distribution of first-year value: ARRA & Non-ARRA contracts (2009-2015)

are more likely to receive medium-sized contracts, whereas lobbying firms are more likely to receive large contracts. Perhaps most striking is the fact that firms that lobby on ARRA account for around 1.5% percent of the total number of firms awarded ARRA contracts. Yet, these firms were awarded 34% of ARRA contracts and 30% of the total ARRA contract spending between 2009 and 2015.



**Notes:** This figure shows distribution of log mean first-year contract value for ARRA contracts awarded to firms that lobbied and did not lobby on ARRA at the firm level between 2009 and 2015. The figure excludes observations with zero contract value.

Figure 2.3: Distribution of average value of ARRA contracts (2009-2015)

### 2.5.2 Pooled OLS Specification

Our first aim is to assess whether lobbying is causally linked to procurement allocation during the stimulus period. We take advantage of detailed disclosures that identify lobbying related to the American Recovery and Reinvestment Act, along with ARRA's transparency provisions, which allow us to identify procurement contracts related to the stimulus package. ARRA contracting began in 2009 and we therefore restrict our analysis to 2009 onward. Our empirical approach is cross-

sectional in nature and compares outcomes of firms that lobbied on ARRA versus those that did not. We estimate the following regression:

$$Y_{it} = \alpha_{st} + \beta LARRA_i + \delta_1 MP3_{it} + \delta_2 NP3_{it} + \delta_3 EMP_{it} + \delta_4 LP3_{it} + \varepsilon_{it}$$
 (2.1)

where  $Y_{it}$  is our outcome variable of interest and  $\alpha_{st}$  captures industry-year fixed effects. MP3 measures the average first-year value of contracts awarded in the previous three years and NP3 measures the total number of new contracts awarded in the previous three years. Both control for the fact that firms may be awarded more and larger ARRA and/or non-ARRA contracts simply because they have experience in handling such contracting volume and size. EMP controls for the correlation between firm size and federal contracting. And LP3 controls for the possibility that corporate lobbying of any kind, rather than targeted lobbying on ARRA, influences outcomes. The coefficient on  $\beta$  captures the effect of lobbying on ARRA (LARRA).

Given our working hypothesis that corporate lobbying influences procurement allocation, we expect the coefficient of  $\beta$  to be positive and significant for outcome variables associated with ARRA contracting, and insignificant for outcome variables not directly associated with ARRA. Our preliminary results from pooled OLS regressions are reported in Table 2.1.

Consistent with our expectations, we find that lobbying on ARRA is positively associated with obtaining an ARRA contract (column 1), the number of ARRA

contracts (column 2) and the total first-year value of these contracts (column 3). However, inconsistent with our expectations, is the fact that lobbying on ARRA is also positively (and significantly) correlated with the number of non-ARRA contracts awarded in the post-2009 period, though it is not significantly correlated with with the value of non-ARRA contracts awarded during this period.

Table 2.1: Corporate lobbying and procurement allocation (pooled OLS regression)

	(1)	(2)	(3)	(4)	(5)
	DARRA	NARRA	VARRA	nonNARRA	nonVARRA
					_
LARRA	$0.0544^{***}$	$0.0844^{***}$	$0.742^{***}$	$0.161^{***}$	0.0108
	(0.00615)	(0.0107)	(0.0826)	(0.0189)	(0.0469)
Ind-Year FE	Yes	Yes	Yes	Yes	Yes
Full Controls	Yes	Yes	Yes	Yes	Yes
Obs.	706,628	706,628	706,628	706,628	706,625
R-sq	0.0827	0.0782	0.0909	0.678	0.378

Notes: The dependent variable in the first column is a dummy whether the firm was awarded any ARRA contract; in the second column is the number of ARRA contracts awarded; and in the third column is the total first-year value of ARRA contracts. The fourth column is the number of non-ARRA contracts awarded and the last column is the total first-year value of non-ARRA contracts. We only report the coefficient of interest, namely the time-invariant dummy LARRA, which is equal to one if the firm lobbies on ARRA. In all regressions we also control for industry-year fixed effects and for firm-level employment, lobbying in the previous three years, and the average value and total number of new contracts awarded in the previous three years. Standard errors are robust. \*\*\*, \*\*, and \* indicates significance at the 1%, 5%, and 10% levels, respectively.

## 2.5.3 Propensity Score Matching Approach

One important concern regarding the full sample used in our analysis thus far is selection bias. As table 2.2 shows, well under 1% of all procurement contractors lobby on ARRA, but those that do are quite different from the rest. In particular, they have earned more and larger contracts in the past (NP3 and MP3); have virtually all lobbied in the previous three years (LP3); and have higher employment

Table 2.2: Pre-matching difference between ARRA lobbying and non-lobbying firms

		Mea		
Variable	Sample	ARRA Lobby	Non Lobby	Bias
#FIRMS	Unmatched	541	117,963	
EMP	Unmatched	2.57	1.30	118.6
LP3	Unmatched	0.99	0.01	877.6
MP3	Unmatched	10.33	7.25	69.8
NP3	Unmatched	3.64	1.51	104.5

**Notes:** The table reports the comparison of means between treated (ARRA lobby) and untreated (Non Lobby) firms in the unmatched sample for all variables that will be used in our propensity score matching.

(EMP). The average ARRA lobbying firm belongs to the size bin representing employment above 50 (and likely closer to above 250) and the average non-lobbying firm belongs to the size bin representing employment below 50. Ideally, we would like to know the counter-factual award of ARRA contracts to firms that lobbied on ARRA if they had instead chosen not to lobby. Since such a counter-factual is unobservable, we are forced to rely on a control group of firms that did not lobby on the stimulus bill. Since lobbying is an endogenous choice and correlated with factors such as firm size, which also affect contract allocation, the sample of all firms that did not lobby on ARRA is not an appropriate counter-factual. We address this selection bias by using a standard propensity score matching approach.

For the results reported in the next section, we restrict ourselves to a sample of firms that lobbied on ARRA in 2009 (treatment group) and a sample that is observationally similar to the treated group, but that did not lobby (control group). In the first stage, we focus on the cross-section of firms in 2009 since this is the year in which treatment status is determined. We estimate a logit model to predict

whether a firm lobbies on ARRA as a function of all the firm characteristics used in the pooled regression, including employment (EMP), industry (NAICS), lobbying status over the previous three years (LP3), number of new contracts awarded in the previous three year (NP3), and the mean value of those contracts (MP3). Formally, we estimate:

$$LARRA_i = \lambda_s + \eta_1 MP3_{it} + \eta_2 NP3_{it} + \eta_3 EMP_{it} + \eta_4 LP3_{it} + \varepsilon_{it}$$
(2.2)

We then use the resulting propensity scores to construct a nearest-neighbor matched sample of firms. Once we identify our control group in 2009, we track the group from that year onward in our second stage regressions. With this approach we eliminate from our sample procurement contractors that are observationally very different from those firms that lobbied on ARRA in 2009.

Before turning to our matched sample estimation results, it's helpful to review the results from our first stage in table 2.3. Because we allow for matching with replacement, our 541 firms that lobbied on ARRA are matched to a sample of 367 firms that did not. As a result, our second stage regressions will employ frequency weight. Table 2.3 confirms that our nearest neighbor matching virtually eliminates selection on observables.

As suggested in Imbens and Rubin (2015), we consider an additional validation exercise. If our matched sample does well in addressing selection, we would expect no differences in pre-ARRA (2006-2008) procurement outcomes for the treated and

Table 2.3: First-stage: post-estimation comparison

		3.5		
		Mea	$\mathbf{n}$	
Variable	Sample	ARRA Lobby	Non Lobby	Bias
#FIRMS	Unmatched	541	117,963	
	Matched	541	367	
EMP	Unmatched	2.57	1.30	118.6
	Matched	2.56	2.52	4.2
LP3	Unmatched	0.99	0.01	877.6
	Matched	0.99	0.99	0.0
MP3	Unmatched	10.33	7.25	69.8
	Matched	10.32	9.96	8.0
NP3	Unmatched	3.64	1.51	104.5
	Matched	3.62	3.47	7.8

**Notes:** The table reports the comparison of means between treated (ARRA Lobby) and untreated (Non Lobby) firms in the unmatched and matched sample for all variables used in the first-stage regressions.

control groups. Table 2.4 considers the total number (NUM) and total first year value (VAL) of contracts in the first two columns. As should be the case, lobbying on ARRA has no significant effect on these outcomes. We might be concerned that these outcomes are highly correlated with the procurement related variables used in our first stage. In column (3) we find that for a non-targeted outcome, the share of contracts awarded competitively (COMP), there is still no statistically significant difference between our treatment and control groups.

The results from our second stage regressions, reported in table 2.5 show that lobbying is indeed influential in shaping the allocation of ARRA-supported procurement contracts. After controlling for selection into lobby by restricting the analysis to a smaller sample of firms that are similar in size, past lobbying and experience in federal procurement, we still find that firms that lobbied on ARRA are on average significantly more likely to win ARRA contracts. Importantly, the magnitudes are

Table 2.4: Placebo Test

	(1) NUM	(2) VAL	(3) COMP
LARRA	-0.0552	-0.1678	0.1906
	(0.039)	(0.108)	(0.0134)
Ind-Year FE	Yes	Yes	Yes
Full Controls	Yes	Yes	Yes
Obs.	839	839	839
Freq. Weighted	Yes	Yes	Yes
R-sq	0.923	0.739	0.117

Notes: The period of analysis is 2006-2008. The dependent variable in the first column total number of contracts; in the second column is total first year value; and in the third column is the share of contracts awarded competitively. We only report the coefficient of interest, namely the time-invariant dummy LARRA, which is equal to one if the firm lobbies on ARRA. In all regressions we also control for industry-year fixed effects and for firm-level employment, lobbying in the previous three years, and the average value and total number of new contracts awarded in the previous three years. Standard errors are robust. \*\*\*, \*\*\*, and \* indicates significance at the 1%, 5%, and 10% levels, respectively.

economically significant. When the regression is evaluated at the mean, the results imply that firms that lobbied on ARRA-related bills won 8% more and 68% larger ARRA-contracts than firms that did not lobby on ARRA.

Although the control group from the matched sample closely resembles the treated group, there is still room for unobserved factors that make firms both lobby more intensively on ARRA and win more contracts. For this reason, we evaluate non-ARRA contract outcomes of ARRA lobbying and non-lobbying firms. In contrast to the pooled regression approach, once we correct for selection, we find that lobbying on ARRA has not impact on the number or size of non-ARRA contracts awarded. Because it is unlikely that the unobserved factors differentially affect ARRA and non-ARRA contracts the relationship uncovered in this section between lobbying and procurement is likely to be causal.

Summarizing, after correcting for selection bias by employing nearest neighbor matching, we find a strong positive effect on the probability, number, and size of ARRA-contracts awarded to firms engaged in corporate lobbying on ARRA and the non-causal correlation between ARRA lobbying and non-ARRA contracting vanishes.

Table 2.5: Corporate lobbying and procurement allocation (Second-stage regression)

	(1)	(2)	(3)	(4)	(5)
	DARRA	NARRA	VARRA	${\rm nonNARRA}$	${\rm nonVARRA}$
LARRA	$0.0469^{***}$	$0.0776^{***}$	$0.679^{***}$	-0.0104	-0.0690
	(0.00802)	(0.0129)	(0.104)	(0.0199)	(0.0531)
Ind-Year FE	Yes	Yes	Yes	Yes	Yes
Full Controls	Yes	Yes	Yes	Yes	Yes
Obs.	5,201	5,201	5,201	5,201	5,201
R-sq	0.240	0.217	0.240	0.899	0.725

Notes: The table reports results for the matched sample of firms obtained from nearest neighbor matching. The dependent variable in the first column is a dummy whether the firm was awarded any ARRA contract; in the second column is the number of ARRA contracts awarded; and in the third column is the total first-year value of ARRA contracts. The fourth column is the number of non-ARRA contracts awarded and the last column is the total first-year value of non-ARRA contracts. We only report the coefficient of interest, namely the time-invariant dummy LARRA, which is equal to one if the firm lobbies on ARRA. In all regressions we also control for industry-year fixed effects and for firm-level employment, lobbying in the previous three years, and the average value and total number of new contracts awarded in the previous three years. Standard errors are robust. \*\*\*, \*\*, and \* indicates significance at the 1%, 5%, and 10% levels, respectively.

# 2.5.4 Matching Approach with Compustat Sample

One threat to the results presented in the previous section is that firms that lobbied on ARRA and won ARRA contracts may simply be more efficient. These firms have the resources to lobby, but win contracts because they have a reputation for being more productive. To address this concern, we focus on a sub-sample

of firms that appear in both the Federal procurement data and Compustat. For these firms we estimate and control for total factor productivity (TFPR) in our regressions.

We begin by augmenting our analysis data with Compustat. Since both the procurement and Compustat data contain firm names and detailed address information, we use an iterative probabilistic name and address matching procedure to combine the two. Through this procedure, we identify nearly 2,100 publicly-listed firms in the procurement data, 300 (14%) of which are awarded ARRA contracts. Of publicly-traded contractors, about 790 (38%) are also present in the lobby data, 25% of which lobby on ARRA. Of the 200 listed firms that lobby on ARRA, about 38% are awarded ARRA contracts.

Following Imrohoroglu and Tuzel (2014), we use firm-level value added, employment, investment, physical capital, and 2-digit NAICS codes to estimate firm-level productivity. We integrate data from the Bureau of Economic Analysis for the GDP deflator, which is used to deflate value added; and the price index for private fixed investment, which is used to deflate investment and capital. TFPR is estimated using the Wooldridge (2009) extension of Levinsohn and Petrin (2003). In the estimation procedure, we choose investment as our proxy for the correlation between inputs and productivity shocks.

The analysis from the previous section is repeated. In particular, we first perform nearest neighbor matching to obtain a sample of treated and control firms that are observationally similar. The first-stage results are reported in table 2.6. Compared to the full Federal procurement sample, as expected, the Compustat

subset consists of firms that are larger, more likely to have lobbied in the previous three years, and more experienced on the procurement market. Comparing now the publicly-listed firms that lobbied on ARRA with those that did not, we observe important differences. Namely, as in the full procurement data, firms that lobbied on ARRA are larger and have more prior experience with lobbying and the procurement market. They are also slightly more productive. These differences suggest that our nearest neighbor matching approach is indeed warranted. Post matching, these observational differences between the two groups are substantially reduced and we proceed to the second stage regressions.

Table 2.6: First-stage: post-estimation comparison

		Mean		
		Mea	UII	
Variable	Sample	ARRA Lobby	Non Lobby	Bias
#FIRMS	Unmatched	143	1,049	
	Matched	143	86	
EMP	Unmatched	3.90	3.56	63.2
	Matched	3.90	3.91	-1.7
LP3	Unmatched	0.98	0.28	210.8
	Matched	0.98	0.98	0.0
MP3	Unmatched	11.67	10.45	50.0
	Matched	11.65	11.86	-8.9
NP3	Unmatched	5.82	3.76	94.5
	Matched	5.63	5.88	-11.3
TFPR	Unmatched	10.04	9.95	8.5
	Matched	10.07	10.32	-21.5

**Notes:** The table reports the comparison of means between treated (ARRA Lobby) and untreated (Non Lobby) firms in the unmatched and matched sample for all variables used in the first-stage regressions.

The results presented in the previous section are confirmed in the publiclylisted subsample. Even after controlling for productivity, we find that lobbying on ARRA influences procurement allocation. Firms that lobby on ARRA are 6% more likely to win ARRA contracts, win 12% more ARRA contracts; and are awarded contracts that are 98% larger than firms that did not lobby on ARRA. Moreover, lobbying on ARRA has no significant effect on non-ARRA procurement outcomes.

Table 2.7: Compustat Sample: Lobbying and procurement allocation (Second-stage regression)

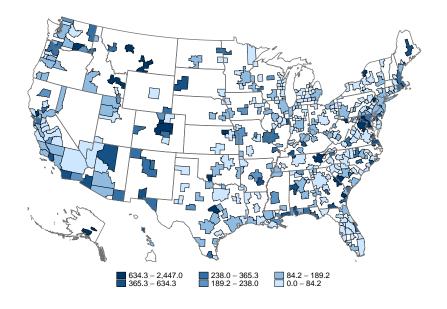
	(1)	(2)	(3)	(4)	(5)
	DARRA	NARRA	VARRA	nonNARRA	nonVARRA
LARRA	0.0592**	0.120***	0.980***	-0.0117	-0.0103
	(0.0247)	(0.0457)	(0.329)	(0.0391)	(0.0654)
ln(TFPR)	-0.00396	0.0132	-0.0235	-0.0744	-0.119
	(0.0215)	(0.0325)	(0.290)	(0.0522)	(0.0797)
Ind-Year FE	Yes	Yes	Yes	Yes	Yes
Full Controls	Yes	Yes	Yes	Yes	Yes
Obs.	1092	1092	1092	1092	1092
R-sq	0.347	0.328	0.385	0.932	0.905

Notes: The table reports results for the matched sample of firms obtained from nearest neighbor matching for the Compustat-Federal procurement subsample. The dependent variable in the first column is a dummy whether the firm was awarded any ARRA contract; in the second column is the number of ARRA contracts awarded; and in the third column is the total first-year value of ARRA contracts. The fourth column is the number of non-ARRA contracts awarded and the last column is the total first-year value of non-ARRA contracts. We only report the coefficients of variables of interest, namely the time-invariant dummy LARRA and estimated productivity TFPR. In all regressions we also control for industry-year fixed effects and for firm-level employment, lobbying in the previous three years, and the average value and total number of new contracts awarded in the previous three years. Standard errors are robust. \*\*\*, \*\*\*, and \* indicates significance at the 1%, 5%, and 10% levels, respectively.

## 2.6 Differential Effects of ARRA Procurement on Employment Growth

# 2.6.1 Descriptive Statistics

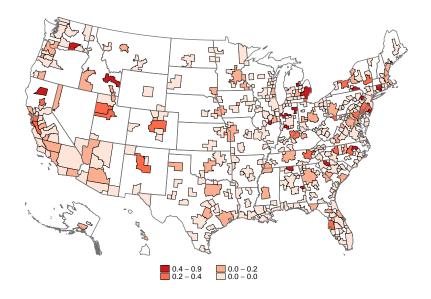
Having shown that lobbying effects the allocation of ARRA-supported contracts, we now turn to whether this influence has implications for local employment outcomes. The foundation of our empirical approach is the well-established use of geographic variation in stimulus spending to identify the effect of this spending on labor market outcomes (Chodorow-Reich, Feiveson, Liscow, and Woolston, 2012; Dube, Kaplan, and Zipperer, 2014; Dupor and Mehkari, 2016; Feyrer, 2011; Wilson, 2012). As figure 2.4 shows, there is quite a bit of variation in ARRA supported federal procurement per capita in 2009 and 2010 across MSAs. This geographic variation helps us achieve identification by asking whether regions that received more money per capita created or saved more jobs.



**Notes:** This figure shows the distribution of total ARRA procurement spending per capita across MSAs between 2009 and 2010.

Figure 2.4: Distribution of ARRA Procurement across MSAs

We are interested in whether ARRA spending channeled through firms that lobbied for ARRA contracts has a different effect on employment than ARRA spending channeled through other firms. To identify this effect, we need sufficient geographic variation to ask whether regions that received less money per capita through lobbying firms were able to create or save more jobs. Figure 2.5 shows the fraction of total ARRA procurement awarded to firms that lobbied on ARRA.



**Notes:** This figure shows the distribution of the share of total ARRA procurement spending channeled through firms that lobbied on ARRA across MSAs between 2009 and 2010.

Figure 2.5: Distribution of Lobbying share of ARRA Procurement across MSAs

### 2.6.2 Baseline OLS Regressions

We are interested in estimating the following baseline model:

$$\frac{L_{i,T} - L_{i,0}}{POP_{i,0}} = \beta_0 + \beta_1 \frac{NonLob_i}{POP_{i,0}} + \beta_2 \frac{Lob_i}{POP_{i,0}} + X'_{i,0}\delta + \gamma + \epsilon_i$$
 (2.3)

The dependent variable is change in employment in MSA i between the first quarter of 2009 and the first quarter of 2011, scaled by the population of the MSA in 2009. We control for a list of covariates  $(X'_{i,0})$  that the literature has found

to strongly predict changes in employment, particularly during the recent financial crisis. These include the size of the economy as measured by the working-age population in 2009 and short-run economic conditions measured by the unemployment rate in 2009. To account for the fact that industry composition is heterogeneous across different MSAs and certain industries are associated with stronger decline in employment to population ratios, we control for the predicted change in employment over 2009-2011 period.<sup>3</sup> To construct the predicted change variable, we first multiply the employment share of sector s in MSA i in 2008 with the national changes in employment in sector s between 2009 and 2011, and sum the products across all s for each MSA. This Bartik-style variable strongly predicts the actual changes in employment to population ratio in MSA i over 2009-2011 period, which allows us to control for the changes in employment caused by industry composition of the area. We also control for the share of young (age 18-24) men with less than a college education to take into account that they exhibited the sharpest decline in the employment-to-population ratio during the crisis. Finally, to account for the fact that the areas experiencing the largest housing boom were also the hardest hit by the financial crisis, we control for the increase in the rate of the housing price index between 2003 and 2007. Finally, we include a dummy variable that takes on a value of one if MSA includes a county represented by a congressperson in a committee that ARRA legislation went through.

Our two variables of interest are  $\frac{NonLob}{POP}$  and  $\frac{Lob}{POP}$ , which measure the total

<sup>&</sup>lt;sup>3</sup>For example, Charles, Hurst, and Notowidigdo (2013) show that the decline in manufacturing industry had a large impact on the decline in non-employment.

amount of ARRA procurement per capita through firms that did not lobby on ARRA and the total amount of ARRA procurement per capita through firms that did lobby, respectively. We are particularly interested in differences between the coefficients  $\beta_1$  and  $\beta_2$ .

We estimate our baseline model first using total ARRA procurement spending per capita. Our OLS results in table 2.8 estimate that every \$1 million in procurement spending between 2009 and 2010 created or saved 6.9 jobs. In the second column, we decompose ARRA procurement spending into that channeled through non-lobbying and lobbying firms. The decomposition suggests that lobbied and non-lobbied contracts have no significant effect on employment growth between 2009 and 2011.

Table 2.8: County Emplyoment Outcomes & ARRA Procurement (OLS)

	(1)	(2)
	$\left. \triangle \left( \frac{EMP}{POP} \right)_{0911} \right.$	
$\frac{ARRA}{POP}$	6.90** (3.07)	
$rac{NonLOB}{POP}$		4.72 $(4.47)$
<u>LOB</u> POP		23.72 (17.46)
State FE	Yes	Yes
Full Controls	Yes	Yes
Observations	338	338

Notes: The table reports results for MSA-level regressions. The dependent variable is the change in employment-to-working age population ratio between the first quarter of 2009 and the first quarter of 2011. We only report the coefficients of variables of interest, namely total ARRA procurement spending per capita  $(\frac{ARRA}{POP})$ , total ARRA spending channeled through firms that did not lobby on ARRA  $(\frac{NonLob}{POP})$ , and total ARRA spending channeled through firms that lobbied on ARRA  $(\frac{LOB}{POP})$ . In all regressions we also control for INSERT LIST OF CONTROLS. Standard errors are robust. \*\*\*, \*\*, and \* indicates significance at the 1%, 5%, and 10% levels, respectively.

One challenge we face that is common to other papers assessing the employment impact of stimulus spending, is that regional variation in spending is not exogenous. During the Great Recession, harder hit areas tended to receive more funding, which biases OLS estimates downwards. We face an additional challenge since we may be missing an unobserved factor at the MSA level that is simultaneously correlated with receiving more procurement through lobbying firms and employment outcomes. For instance, if a firm is a large employer in the MSA, it will have a large effect on employment. And if at the same time, that firm is being adversely affected in the recession, it may lobby on the stimulus package to insulate its sales or employment from the crisis as much as possible. In the next section we attempt to address these concerns by constructing a set of Bartik instruments to use in our estimation.

#### 2.6.3 Instrumental Variables

Let us first consider the concern that harder hit areas receive more funding. Others have addressed this issue by using various instrumentation strategies, including the use of exogenous formulary allocation factors Wilson (2012) and pre-recession spending levels Chodorow-Reich, Feiveson, Liscow, and Woolston (2012). Our approach is similar in spirit to Chodorow-Reich, Feiveson, Liscow, and Woolston (2012) who use state pre-recession Medicaid spending levels to instrument for ARRA state fiscal relief.

We take full advantage of our detailed federal procurement data, which covers

the period 2005 onward. Our instrument for total ARRA procurement is a Bartik measure that relies on variation in national ARRA spending across sectors, and on variation across sectors in the degree of reliance on different counties prior to the recession. To be more precise, we construct a predicted measure of total ARRA procurement to a MSA i ( $\bar{A}RR\bar{A}_i$ ) by first calculating total U.S. ARRA procurement channeled to each three-digit NAICs sector after excluding procurement to a MSA i. This gives us a measure of national allocation to each sector that is exogenous of the particular MSA. We then calculate the fraction of overall U.S. procurement in each sector that was channeled to MSA i in 2007. This measure of reliance, since it is measured prior to the passage of ARRA, should be unaffected by the stimulus package. We use these two measures to allocate ARRA spending in each sector to MSA i and obtain a MSA-level measure by aggregating across all sectors. The strength of our instrument relies on the validity of two assumptions: that i) if the government is heavily reliant on a particular MSA for procurement in sector s, it is likely to remain reliant during the stimulus period; and ii) if a large amount of ARRA is channeled to sector s, then counties on which the government was previously more reliant might be expected to get a greater fraction of that spending.

$$\widehat{ARRA}_i = \sum_{s} ProcShare_{\{i,s\}} ARRA_{sn\{i,s\}}$$
 (2.4)

Once we instrument for total ARRA spending, we are still faced with the potential endogeneity of ARRA spending channeled through lobbying firms. To address this, we take advantage of the fact that corporate lobbying is persistent Kerr,

Lincoln, and Mishra (2014a) and that there is substantial variation across sectors in the fraction of procurement awarded through lobbying firms. Since lobbying is associated with high fixed costs, it is not surprising that over 90\% of firms that lobbied on ARRA in 2009 had previously lobbied on other issues. Further, some sectors channel as little as 1% of procurement through lobbying firms, while others channel as much as 80%. The strength of our instrument relies on the assumption that sectors that heavily relied on lobbying firms prior to the crisis will continue to do so, and that these lobbying firms are likely to continue to lobby in 2009. In particular, to construct our instrument  $(\widehat{LOB}_i)$  for MSA i we first calculate the share of total procurement spending channeled through lobbying firms for each sector in 2007. We then combine this share with our how reliant procurement in each sector was on MSA i in 2007 and how much ARRA spending, net of the amount to MSA i, went to each sector. Our MSA-level instrument is obtained by aggregating across all sectors. Finally, we construct our instrument for procurement through nonlobbying firms by taking the difference between predicted total ARRA procurement and predicted ARRA procurement through lobbying firms.

$$\widehat{LOB}_i = \sum_{s} ProcShare_{\{i,s\}} \times ARRA_{sn\{i,s\}} \times Lob_s$$
 (2.5)

$$\widehat{NonLOB}_i = \widehat{ARRA}_i - \widehat{LOB}_i \tag{2.6}$$

In table 2.9 we first report the results from the first stage of our 2SLS regression model. We include, but do not report, the full set of controls used in our

baseline regressions. The results confirm that our instruments are strong predictors of actual ARRA spending, as well as ARRA spending through lobbying and non-lobbying firms. As expected, the relationship between all of the actual and the predicted measures is positive and significant.

Table 2.9: First-Stage Instrumental Variable Regression Results

$\frac{ARRA}{POP}$	$\frac{NonLOB}{POP}$	$\frac{LOB}{POP}$
0.616*** (0.14)		
	0.554*** (0.12)	-0.404 (0.50)
	-0.027 $(0.02)$	0.466** (0.19)
YES	YES	YES
YES	YES	YES
338	338	338
0.35	0.39	0.31
	POP 0.616*** (0.14) YES YES YES 338	POP POP  0.616*** (0.14)  0.554*** (0.12)  -0.027 (0.02)  YES YES YES YES YES 338 338

p < 0.10, \*\* p < 0.05, \*\*\* p < 0.01

Notes: The table reports results of the first stage for our IVs. The dependent variable in the first column is total ARRA procurement spending per capita  $(\frac{ARRA}{POP})$ ; in the second column is total ARRA spending channeled through firms that did not lobby on ARRA  $(\frac{NonLob}{POP})$ ; and in the third column is total ARRA spending channeled through firms that lobbied on ARRA  $(\frac{LOB}{POP})$ . We only report the coefficient estimates for our Bartik instruments  $(\frac{ARRA}{POP}, \frac{NonLob}{POP}, \frac{LOB}{POP}, \frac{LOB}{POP})$ , respectively. In all regressions we also control for INSERT LIST OF CONTROLS Standard errors are robust.

\*\*\*, \*\*, and \* indicates significance at the 1%, 5%, and 10% levels, respectively.

## 2.6.4 Regression Results

Instrumenting for ARRA procurement spending, we estimate the effect of the stimulus on net employment growth in table 2.10. In the first column we estimate that every \$1 million in ARRA supported procurement spending saves or creates

11.5 jobs. Our estimate is similar to the estimate of 8 jobs saved reported in Dube, Kaplan, and Zipperer (2014) and Wilson (2012).

In the second column of table 2.10 we find a striking result. The entire employment effect of ARRA procurement spending is explained by money channeled through non-lobbying firms. Every \$1 million in ARRA-supported procurement through these firms supported 16 jobs. In contrast, the employment effect of procurement through firms that lobbied on ARRA is insignificant. Our results suggest that the impact of stimulus spending weakened when procurement is allocated to firms that lobby. Note that the employment multiplier is much two times larger when we move from the first column to the coefficient for non-lobbying contractors on the second column. This large change comes from the fact that a significant share of procurement money is going to lobbying firms. In short, we find that it is not only the amount of stimulus spending that matters for employment outcomes, but also how that spending is allocated.

Table 2.10: MSA Employment Outcomes & ARRA Procurement (IV regression)

	$\Delta \left(\frac{EMP}{POP}\right)_{(11-09)}$	$\Delta \left(\frac{EMP}{POP}\right)_{(11-09)}$
$\frac{\widehat{ARRA}}{POP}$	11.48*** (4.35)	
$\frac{\widehat{NonLob}}{POP}$		16.07** (6.49)
LOB POP		-1.28 (35.79)
Census Division FE	Yes	Yes
Full Controls	Yes	Yes
Observations	338	338
F-statistic	141.7	40.5
* n < 0.10 ** n < 0.05 *	*** n < 0.01	

<sup>\*</sup> p < 0.10, \*\* p < 0.05, \*\*\* p < 0.01

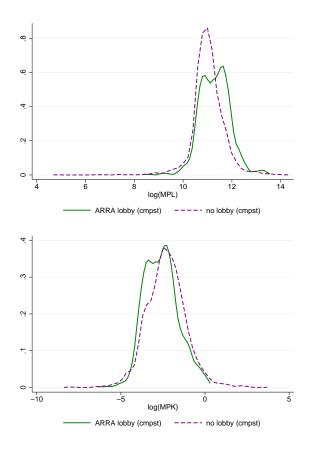
Notes: The table reports results for MSA-level regressions after instrumentation for variables of interest. The dependent variable is the change in employment-to-working age population ratio between the first quarter of 2009 and the first quarter of 2011. We only report the coefficients of variables of interest, namely the Bartik instruments for total ARRA procurement spending per capita  $(\widehat{\frac{ARRA}{POP}})$ , total ARRA spending channeled through firms that did not lobby on ARRA  $(\widehat{\frac{NonLob}{POP}})$ , and total ARRA spending channeled through firms that lobbied on ARRA  $(\widehat{\frac{LOB}{POP}})$ . In all regressions we also control INSERT LIST OF CONTROLS. Standard errors are robust. \*\*\*, \*\*\*, and \* indicates significance at the 1%, 5%, and 10% levels, respectively.

# 2.7 A Simple Theoretical Explanation

In this section, we provide a simple theoretical explanation for why we find a smaller employment growth effect from procurement spending that is channeled through lobbying firms. The intuition is as follows. Suppose firms lobby in period t and procurement contracts are awarded in period t + 1. As shown in Section 2.5, lobbying firms win government contracts with a higher probability. Therefore, if private demand is equal, lobbying firms expect a higher future demand compared to the firms that do not lobby. Higher expected demand incentivizes lobbying firms

to accumulate more physical capital in t. If lobbying firms indeed win government contracts in t+1, they can fulfill the demand with less labor input, and fewer jobs are created for each procurement dollar spent in t+1.

This simple explanation can tested using our matched Federal procurement-Compustat sample. In particular, if the posited theoretical explanation is true, we should observe that firms lobbying on ARRA have higher marginal product of labor (MPL) and lower marginal product of capital (MPK) than firms that do not lobby on ARRA. Figure 2.6, tests this hypothesis by comparing the distributions of log(MPL) and log(MPK) for firms that lobbied on ARRA versus those that did not. These figures suggest that firms that lobby on ARRA have higher MPL. The results for MPK are less clear, though it appears that lobbying firms do have slightly higher MPK than those that do not.



**Notes:** This figure shows the distribution of log(marginal product of labor) [left panel] and log(marginal product of capital) [right panel], separately for firms that lobby on ARRA (solid green line) and firms that do not (dashed purple line). The underlying sample is the Federal procurement data matched to Compustat.

Figure 2.6: Distribution of MPL and MPK: Lobbying vs. Non-lobbying firms

We now test this hypothesis more formally. We first match firms that lobbied on ARRA (treatment) with those that did not (control) based on employment, past lobbying experience, past procurement experience, and industry. We then take the resulting matched sample and regress LARRA on the same variables, as well as log(MPL) and log(MPK). The results in column 2 of table 2.11 are consistent with our hypothesis. Firms lobbying on ARRA have higher MPL and lower MPK, even after matching on other firm-level characteristics. As our theoretical explanation

suggests, it could be that lobbying firms adopt capital intensive technologies and therefore do not hire more labor.

Table 2.11: Compustat Sample MPL and MPK Regression Results)

	(1)	(2)
$\overline{ln(MPL)}$		0.174***
		(0.0277)
ln(MPK)		-0.0721***
		(0.0157)
NP3	$0.0214^{**}$	0.0313***
	(0.00880)	(0.00840)
MP3	0.0283***	0.0286***
	(0.00964)	(0.00916)
LP3	0.572***	0.580***
	(0.0277)	(0.0289)
Ind-Year FE	Yes	Yes
Size bins	Yes	Yes
Obs.	1,176	1,117
R-sq	0.210	0.259

Notes: The table reports results for firm-level regressions post implementation of nearest neighbor matching. The dependent variable is a time-invariant dummy LARRA, which is equal to one if the firm lobbies on ARRA. The main coefficients of interest are those associated with ln(MPL) and ln(MPK). In all regressions we also control for industry-year fixed effects and for firm-level employment, lobbying in the previous three years, and the average value and total number of new contracts awarded in the previous three years. Standard errors are robust. \*\*\*, \*\*, and \* indicates significance at the 1%, 5%, and 10% levels, respectively.

### 2.8 Conclusion

This paper studies the allocation and effectiveness of procurement spending during the largest stimulus package in American history. In particular, by matching data on firm lobbying and federal procurement contracts during the American Reinvestment and Recovery Act of 2009, we investigate i) how firm lobbying activities impact the allocation of procurement spending across firms, and ii) if the effect

on MSA employment differs when most of the contracts are channeled by lobbying contractors.

We answer the first question using a propensity score matching approach and estimate that firms that lobbied on ARRA win 5.3% more and 50% larger ARRA-supported contracts. We validate our identification strategy by showing that after matching, ARRA lobbying only influences ARRA procurement outcomes and does not have any significant impact on past or future corresponding non-ARRA contract outcomes.

We tackle the second question using Bartik instruments to address potential endogeneity concerns. We estimate that \$1 million of total procurement spending yields 11.5 jobs at the MSA level, which is broadly consistent with previous estimates in the literature. Yet, we find that this effect is entirely driven by non-lobbying firms; \$1 million channeled through non-lobbying firms yields 16 jobs, while that channeled through lobbying firms has no significant effect on employment.

We also provide a simple theoretical explanation on why procurement spending on lobbying firms generated fewer jobs. Lobbying firms in general tend to expect to win government contracts with higher probability. Facing higher expected demand from the public sector, lobbying firms tend to accumulate more physical capital in advance. Therefore, when lobbying firms indeed win government contracts, they can fulfil the demand with fewer workers. We document indirect evidence for this explanation by showing that lobbying firms tend to exhibit lower MPK and higher MPL compared to non-lobbying firms, even after controlling for size and industry. Our explanation is not necessarily based on rent-seeking behavior or distortions.

However, our findings cautions that when the allocation of government spending is affected by firms' lobbying behavior, job creation effect may be mitigated.

### Appendices

# A Appendix for Chapter 1

### A.1 Simple Model Solution Derivation

Note that the entrepreneur in period 1 chooses labor  $n_1$  to maximize period 1 profits and it does not affect period 2 expected utility. Explicitly writing out  $V_2(z_1, h)$ , the optimization problem of  $\Delta$  can be re-written as

$$\max_{\Delta \geq 0} \ln(\Gamma z_1) + \left\{ e^{-\gamma \Delta} \cdot \max\{\ln(\Gamma z_1 e^{\Delta}), \ln(wh)\} + (1 - e^{-\gamma \Delta}) \cdot \max\{\ln(\Gamma z_1 e^{-\Delta}), \ln(wh)\} \right\}$$

There are four possible objective functions depending on the realization of  $z_2$  and the occupational choice. First, he can stay in business regardless of the  $z_2$  realization, which delivers

$$V_1^{E,E}(z_1,h) = \max_{\Delta \ge 0} \ln(\Gamma z_1) + \beta \left\{ e^{-\gamma \Delta} \cdot \ln(\Gamma z_1 e^{\Delta}) + (1 - e^{-\gamma \Delta}) \cdot \ln(\Gamma z_1 e^{-\Delta}) \right\}$$
(7)

Second, he can stay in business in the high  $z_2$  outcome and exit in the low  $z_2$  outcome. In this case he gets

$$V_1^{E,W}(z_1,h) = \max_{\Delta \ge 0} \ln(\Gamma z_1) + \beta \left\{ e^{-\gamma \Delta} \cdot \ln(\Gamma z_1 e^{\Delta}) + (1 - e^{-\gamma \Delta}) \cdot \ln(wh) \right\}$$
(8)

Third, he can exit regardless of the  $z_2$  realization, in which case he gets the value

$$V_1^{W,W}(z_1, h) = \max_{\Delta \ge 0} \ln(\Gamma z_1) + \beta \ln(wh)$$
 (9)

Lastly, he can stay in business in the low  $z_2$  and exit in the high  $z_2$ .

$$V_1^{W,E}(z_1,h) = \max_{\Delta \ge 0} \ln(\Gamma z_1) + \beta \left\{ e^{-\gamma \Delta} \cdot \ln(wh) + (1 - e^{-\gamma \Delta}) \cdot \ln(\Gamma z_1 e^{\Delta}) \right\}$$
 (10)

One can solve the problem by first deriving the optimal  $\Delta^*$  and the associated value functions conditioning on each case, and then finding the upper envelope of the conditional value functions over  $(z_1, h)$ .

Case 1: 
$$V_1^{E,E}(z_1, h)$$

The Kuhn-Tucker theorem implies that the necessary conditions for the optimal  $\Delta^*$  are

$$-2\Delta^* \gamma e^{-\gamma \Delta^*} + 2e^{-\gamma \Delta^*} - 1 \le 0$$
$$\left(-2\Delta^* \gamma e^{-\gamma \Delta^*} + 2e^{-\gamma \Delta^*} - 1\right) \cdot \Delta^* = 0$$
$$\Delta^* \ge 0$$

Since  $\Delta^* = 0$  violates the first condition,  $\Delta^*$  is strictly positive. Thus  $\Delta^*$  is the root of  $-2\Delta^*\gamma e^{-\gamma\Delta^*} + 2e^{-\gamma\Delta^*} - 1 = 0$ . Denote the solution as  $\bar{\Delta}(\gamma)$ . The implicit function theorem implies that  $\bar{\Delta}$  is decreasing in  $\gamma$ .

Note that equation (7) can be re-written as

$$V^{E,E}(z_1,h) = \ln(\Gamma z_1) + \beta \left\{ \ln(\Gamma z_1) + 2\bar{\Delta}e^{-\gamma\bar{\Delta}} - \bar{\Delta} \right\}$$

Define the second term inside of the bracket as  $C(\gamma) = 2\bar{\Delta}e^{-\gamma\bar{\Delta}} - \bar{\Delta}$ .  $C(\gamma)$  is non-negative at the optimum. It is because if  $C(\gamma)$  were negative, a higher objective function value can be achieved under  $\Delta^* = 0$ , which violates the first necessary condition.

Case 2: 
$$V_1^{E,W}(z_1, h)$$

The necessary conditions for the optimal  $\Delta^*$  are

$$-\ln(\Gamma z_1) - \Delta^* + \frac{1}{\gamma} + \ln(wh) \le 0$$
$$\left(-\ln(\Gamma z_1) - \Delta^* + \frac{1}{\gamma} + \ln(wh)\right) \cdot \Delta^* = 0$$
$$\Delta^* \ge 0$$

Define  $\bar{h}(z_1) = \frac{e^{\frac{-1}{\gamma}}\Gamma z_1}{w}$ . Note that  $-\ln(\Gamma z_1) + \frac{1}{\gamma} + \ln(w\bar{h}(z_1)) = 0$ . Then the optimal  $\Delta^*$  can be characterized as

$$\Delta^* = \begin{cases} \ln(wh) - \ln(\Gamma z_1) + \frac{1}{\gamma} & h \ge \bar{h}(z_1) \\ 0 & h < \bar{h}(z_1) \end{cases}$$

Replacing  $\Delta$  in equation (8) with  $\Delta^*$ , the conditional value function is solved as

$$V^{E,W}(z_1, h) = \begin{cases} \ln(\Gamma z_1) + \beta \left\{ \ln(wh) + \frac{1}{\gamma} (\frac{\Gamma z_1}{wh})^{\gamma} e^{-1} \right\} & h \ge \bar{h}(z_1) \\ \ln(\Gamma z_1) + \beta \ln(\Gamma z_1) & 0 \le h < \bar{h}(z_1) \end{cases}$$

Case 3: 
$$V_1^{W,W}(z_1, h)$$

This case can be ignored as it is strictly dominated by  $V^{E,W}$ . Suppose  $wh \geq \Gamma z_1$ . Then  $h > \bar{h}$  and thus  $V^{E,W} > V^{W,W}$ . On the other hand, if  $wh < \Gamma z_1$ ,  $V^{E,W} > V^{W,W}$  regardless of the value of h within the range.

Case 4: 
$$V_1^{W,E}(z_1, h)$$

This case can be ignored as it is weakly dominated by  $V^{W,W}$  and  $V^{E,E}$ . Suppose  $wh \geq \Gamma z_1$ . Then  $\ln(wh) - \ln(\Gamma z_1) \geq 0$ . Subtracting  $V^{W,E}$  from  $V^{W,W}$ , one can obtain  $V^{W,E} - V^{W,W} = (1 - e^{-\gamma \Delta}) \left( \ln(wh) - \ln(\Gamma z_1) + \Delta \right)$ , which is weakly positive. On the other hand, suppose  $wh < \Gamma z_1$ . Then  $\ln(\Gamma z_1) > \ln(wh)$ . Subtracting  $V^{W,E}$  from  $V^{E,E}$ , one can obtain  $V^{E,E} - V^{W,E} = e^{-\gamma \Delta} \left( \ln(\Gamma z_1) + \Delta - \ln(wh) \right)$ , which is strictly positive.

Therefore it only requires comparing  $V_1^{E,E}(z_1,h)$  and  $V_1^{E,W}(z_1,h)$  to uncover the upper envelope and the optimal solution  $\Delta^*$ . This can be done by fixing  $z_1$  to an arbitrary value and varying the value of h. First, consider h=0. At this point,  $V^{E,W}(z_1,0)=\ln(\Gamma z_1)+\beta\ln(\Gamma z_1)$  and  $V^{E,E}(z_1,0)=\ln(\Gamma z_1)+\beta(\ln(\Gamma z_1)+C(\gamma))$ . Since  $C(\gamma)>0$ ,  $V^{E,W}(z_1,0)< V^{E,E}(z_1,0)$ .

Note that  $\frac{\partial V^{E,E}(z_1,h)}{\partial h} = 0$ , thus  $V^{E,E}$  stays constant for all values of h. On the

other hand,  $\frac{\partial V^{E,W}(z_1,h)}{\partial h} = 0$  for  $0 \le h < \bar{h}(z_1)$ , and  $\frac{\partial V^{E,W}(z_1,h)}{\partial h} > 0$  for all  $h > \bar{h}(z_1)$ . Therefore as h moves from 0 to infinity,  $V^{E,W}$  continuously increases starting from  $\bar{h}$ , and  $V^{E,W}$  and  $V^{E,E}$  crosses once and only once at a value  $h^*(z_1) > \bar{h}(z_1)$ . Therefore, the optimal  $\Delta^*$  can be characterized by

$$\Delta^* = \begin{cases} \ln(wh) - \ln(\Gamma z_1) + \frac{1}{\gamma} & h \ge h^*(z_1) \\ \bar{\Delta}(\gamma) & 0 \le h < h^*(z_1) \end{cases}$$

### A.2 Robustness Check: Sole-proprietor Sample

This section shows that the empirical evidence presented in section 1.3 are robust to restricting the sample to sole-proprietor firms. The purpose of this robustness analysis is to show an evidence that the results are not likely driven by potential errors in the founder approximation method. Table A1 reports the results from the linear probability regression where the dependent variable is the firm exit indicator. Table A2 reports the regression on firm-level growth dispersion, and Table A3 shows results for growth conditioning on survival. In contrast to the main regressions, I find insignificant coefficient for labor productivity growth while all other results are robust. Lastly, Table A4 shows results for the Hurst-Pugsley sector indicator interactions which shows less consistency for sole-proprietor firms.

Table A1: Firm Exit Regressions for Sole-Proprietors

	(1)	(2)	(3)	(4)
	Exit	Exit	Exit	Exit
Log prior earnings	-0.015***	0.001	0.015***	0.016***
	(0.005)	(0.003)	(0.002)	(0.002)
Lagged log employment		-0.060***	-0.112***	-0.113***
		(0.002)	(0.003)	(0.004)
Lagged log labor prod.			-0.132***	-0.135***
			(0.007)	(0.008)
Lagged log wage				0.004
				(0.003)
Founder average age				-0.002***
				(0.000)
Founder male share				0.005
				(0.003)
Ind-Year FE	No	Yes	Yes	Yes
Firm age FE	No	Yes	Yes	Yes
State FE	No	No	No	Yes
Birth year FE	No	No	No	Yes
Obs.	450000	450000	450000	450000
R-sq	0.001	0.104	0.194	0.197

**Notes:** The table reports results for a linear probability regression where the dependent variable is firm exit indicator. The sample is restricted to sole-proprietors, whose business ownership information can be obtained from the Business Register (BR). Standard errors are clustered at the industry (NAICS4) level. \*\*\*, \*\*, and \* indicate significance at the 1%, 5%, and 10% level, respectively.

Table A2: Growth Dispersion Regressions for Sole-Proprietors

	(1)	(2)	(3)	(4)
	$\epsilon^2 \text{ (Rev)}$	$\epsilon^2$ (Rev)	$\epsilon^2 \text{ (Prod)}$	$\epsilon^2 \text{ (Prod)}$
Log prior earnings	0.008***	0.007***	0.005***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)
Lagged log employment		0.002***		-0.003***
		(0.000)		(0.000)
Ind-Year FE	No	Yes	No	Yes
Firm age FE	No	Yes	No	Yes
Obs.				
R-sq	360000	360000	360000	360000
r2	0.001	0.023	0.000	0.021

**Notes:** The table reports results from estimating equation (1.7) for sole-proprietor sample.  $\epsilon_{ijt}^2$  ( $\Delta$  Rev) and  $\epsilon_{ijt}^2$  ( $\Delta$  Prod) are the squared deviations obtained from equation (1.6) where  $Y_{ijt}$  are revenue and labor productivity, respectively. \*\*\*, \*\*, and \* indicate significance at the 1%, 5%, and 10% level, respectively.

Table A3: Growth Regressions for Sole-Proprietor Continuers

	(1)	(2)	(3)
	$\Delta \text{ Rev}$	$\Delta \operatorname{Prod}$	$\Delta \ \mathrm{Emp}$
Log prior earnings	0.009***	-0.003	0.012***
	(0.003)	(0.005)	(0.003)
Lagged log labor prod.	-0.078***	-0.198***	0.119***
	(0.006)	(0.007)	(0.007)
Lagged log employment	-0.049***	0.063***	-0.116***
	(0.004)	(0.006)	(0.008)
Lagged log wage	0.020***	-0.009***	0.032***
	(0.003)	(0.003)	(0.003)
Founder average age	-0.002***	-0.000***	-0.002***
	(0.000)	(0.000)	(0.000)
Founder male share	0.025***	0.0434***	-0.019***
	(0.003)	(0.005)	(0.004)
Ind-Year FE	Yes	Yes	Yes
Firm age FE	Yes	Yes	Yes
State FE	Yes	Yes	Yes
Birth year FE	Yes	Yes	Yes
Obs.	360000	360000	360000
R-sq	0.063	0.139	0.155

**Notes:** The table reports results for OLS regression of firm growth on prior earnings where the sample is restricted to the sole-proprietors. All growth measures are calculated as the DHS growth rate. Standard errors are clustered at the industry (NAICS4) level. \*\*\*, \*\*, and \* indicate significance at the 1%, 5%, and 10% level, respectively.

Table A4: Regression Results with the Hurst-Pugsley Sector Interactions: Sole-Proprietor Sample

	(1)	(2)	(3)	(4)	(5)
	Exit	$\Delta \stackrel{\cdot}{\mathrm{Rev}}$	$\Delta$ Prod	$\Delta \ \mathrm{Emp}$	$\epsilon^2 \; (\text{Rev})$
Log prior earnings	0.014***	0.009***	-0.003	0.011***	0.008***
	(0.001)	(0.002)	(0.004)	(0.003)	(0.001)
$HP \times log prior earnings$	0.003	0.000	-0.001	0.001	-0.002***
	(0.002)	(0.004)	(0.008)	(0.005)	(0.001)
Lagged log labor prod.	-0.135***	-0.078***	-0.198***	0.119***	
208804 108 10001 Prod.	(0.008)	(0.006)	(0.007)	(0.007)	
	(0.000)	(0.000)	(0.001)	(0.001)	
Lagged log employment	-0.113***	-0.049***	0.063***	-0.117***	0.002***
00 0 1	(0.004)	(0.004)	(0.006)	(0.008)	(0.000)
	,	,	,	, , ,	, ,
Lagged log wage	0.004	0.020***	-0.009***	0.032***	
	(0.003)	(0.003)	(0.003)	(0.003)	
Founder average age	-0.002***	-0.002***	-0.000**	-0.002***	
rounder average age	(0.002)	(0.002)	(0.000)	(0.002)	
	(0.000)	(0.000)	(0.000)	(0.000)	
Founder male share	0.005**	0.025***	0.043***	-0.019***	
	(0.003)	(0.003)	(0.005)	(0.004)	
T 117 DD	***	3.7	3.7	3.7	
Ind-Year FE	Yes	Yes	Yes	Yes	Yes
Firm age FE	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	No
Birth year FE	Yes	Yes	Yes	Yes	No
Obs.	450000	360000	360000	360000	360000
R-sq	0.197	0.063	0.139	0.155	0.023

**Notes:** The table reports results for linear regressions re-estimated after including the interaction between HP indicator and log prior earnings. The sample only includes sole-proprietor firms. Standard errors are clustered at the industry level (NAICS4). \*\*\*, \*\*\*, and \* indicate significance at the 1%, 5%, and 10% level, respectively.

### Bibliography

- Abowd, John M., Bryce E. Stephens, Lars Vilhuber, Fredrik Andersson, Kevin L. McKinney, Marc Roemer, and Simon Woodcock (2009), "The LEHD infrastructure files and the creation of the quarterly workforce indicators." In *Producer dynamics: New Evidence from Micro Data*, 149–230, University of Chicago Press.
- Acemoglu, Daron, Ufuk Akcigit, Nicholas Bloom, and William R Kerr (2013), "Innovation, reallocation and growth." NBER Working Paper #18993.
- Acemoglu, Daron and Dan Cao (2015), "Innovation by entrants and incumbents." Journal of Economic Theory, 157, 255–294.
- Adelino, Manuel and Serdar Dinc (2014), "Corporate distress and lobbying: evidence from the stimulus act." *Journal of Financial Economics*, 114, 256–272.
- Aghion, Philippe and Peter Howitt (1992), "A model of growth through creative destruction." *Econometrica*, 60, 323–351.
- Akcigit, Ufuk, Harun Alp, and Michael Peters (2016), "Lack of selection and limits to delegation: firm dynamics in developing countries." NBER Working Paper #21905.
- Akcigit, Ufuk and William R. Kerr (forthcoming), "Growth through heterogeneous innovations." *Journal of Political Economy*.
- Alon, Titan, David Berger, Robert Dent, and Benjamin Pugsley (2017), "Older and slower: The startup deficits lasting effects on aggregate productivity growth." NBER Working Paper #23875.
- Arayavechkit, Tanida, Felipe Saffie, and Minchul Shin (2014), "Capital-based corporate tax benefits: endogenous misallocation through lobbying." Working Paper.
- Arkolakis, Costas, Theodore Papageorgiou, and Olga A. Timoshenko (2017), "Firm learning and growth." Review of Economic Dynamics.
- Atkeson, Andrew and Ariel Tomas Burstein (2010), "Innovation, firm dynamics, and international trade." *Journal of Political Economy*, 118, 433–484.

- Atkeson, Andrew and Patrick J Kehoe (2007), "Modeling the transition to a new economy: Lessons from two technological revolutions." *American Economic Review*, 97, 64–88.
- Bajari, Patrick, Robert McMillan, and Steven Tadelis (2008), "Auctions versus negotiations in procurement: an empirical analysis." *The Journal of Law, Economics and Organization*, 25, 372–399.
- Bajari, Patrick and Steven Tadelis (2001), "Incentives versus transaction costs: a theory of procurement contracts." *The RAND Journal of Economics*, 32, 387–407.
- Baptista, Rui, Francisco Lima, and Miguel Torres Preto (2012), "How former business owners fare in the labor market? Job assignment and earnings." *European Economic Review*, 56, 263–276.
- Bhattacharya, Vivek and Andrew Sweeting (2015), "Selective entry and auction design." *International Journal of Industrial Organization*, 43.
- Bianchi, Milo and Matteo Bobba (2012), "Liquidity, risk, and occupational choices." *Review of Economic Studies*, 80, 491–511.
- Blau, Benjamin, Tyler Brough, and Diana Thomas (2013), "Corporate lobbying, political connections and the bailout of banks." *Journal of Banking and Finance*, 37, 3007–3017.
- Bloom, Nicholas (2009), "The impact of uncertainty shocks." *Econometrica*, 77, 623–685.
- Bombardini, Matilde (2008), "Firm heterogeneity and lobby participation." *Journal of International Economics*, 75, 329–348.
- Bombardini, Matilde and Francesco Trebbi (2012), "Competition and political organization: Together or alone in lobbying for trade policy?" *Journal of International Economics*, 87, 18–26.
- Brogaard, Jonathan, Matthew Denes, and Ran Duchin (2016), "Political influence and government investment: Evidence from contract-level data." Working Paper.
- Bruce, Donald and Herbert J. Schuetze (2004), "The labor market consequences of experience in self-employment." *Labour Economics*, 11, 575–598.
- Buera, Francisco and Juan Pablo Nicolini (2017), "Liquidity traps and monetary policy: Managing a credit crunch." Working paper.
- Buera, Francisco J. and Roberto N. Fattal-Jaef (2016), "The dynamics of development: Innovation and reallocation." Working paper.
- Buera, Francisco J., Joseph P. Kaboski, and Yongseok Shin (2011), "Finance and development: A tale of two sectors." *American Economic Review*, 101, 1964–2002.

- Buera, Francisco J. and Yongseok Shin (2013), "Financial frictions and the persistence of history: A quantitative exploration." *Journal of Political Economy*, 121, 221–272.
- Caggese, Andrea (2016), "Financing constraints, radical versus incremental innovation, and aggregate productivity." Working paper.
- Castro, Rui, Gian Luca Clementi, and Glenn MacDonald (2009), "Legal institutions, sectoral heterogeneity, and economic development." The Review of Economic Studies, 76, 529–561.
- Charles, Kerwin, Erik Hurst, and Matthew Notowidigdo (2013), "Housing booms, manufacturing declines, and labor market outcomes." Working Paper 18949, National Bureau of Economic Research.
- Chodorow-Reich, Gabriel, Laura Feiveson, Zachary Liscow, and William Woolston (2012), "Does state fiscal relief during recessions increase employment? evidence from the american recovery and reinvestment act." *American Economic Journal: Economic Policy*, 4, 118–145.
- Davis, Steven J. and John Haltiwanger (2014), "Labor market fluidity and economic performance." NBER Working Paper #20479.
- Davis, Steven J., John Haltiwanger, and Scott Schuh (1996), Job Creation and Destruction. Cambridge, MA: MIT Press.
- Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda (2014), "The role of entrepreneurship in us job creation and economic dynamism." *Journal of Economic Perspectives*, 28, 3–24.
- Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda (2016), "Where has all the skewness gone? the decline in high-growth (young) firms in the US." *European Economic Review*, 86, 4–23.
- Dillon, Eleanor W. and Christopher T. Stanton (2017), "Self-employment dynamics and the returns to entrepreneurship." NBER Working Paper #23168.
- Dube, Arindrajit, Ethan Kaplan, and Ben Zipperer (2014), "Excess capacity and heterogeneity in the fiscal multiplier: Evidence from the obama stimulus package." Working Paper.
- Duchin, Ran and Denis Sosyura (2012), "The politics of government investment." Journal of Financial Economics, 106, 24–48.
- Dunne, Timothy, Mark J Roberts, and Larry Samuelson (1989), "The growth and failure of us manufacturing plants." *The Quarterly Journal of Economics*, 104, 671–698.

- Dupor, Bill and M. Saif Mehkari (2016), "The 2009 recovery act: Stimulus at the extensive and intensive labor margins." *European Economic Review*, June, 208–228.
- Evans, David S (1987), "The relationship between firm growth, size, and age: Estimates for 100 manufacturing industries." The journal of industrial economics, 567–581.
- Evans, David S. and Boyan Jovanovic (1989), "An estimated model of entrepreneurial choice under liquidity constraints." *Journal of Political Economy*, 97, 808–27.
- Fairlie, Robert W. and Javier Miranda (2017), "Taking the leap: The determinants of entrepreneurs hiring their first employee." *Journal of Economics & Management Strategy*, 26, 3–34.
- Feyrer, Bruce, James and Sacerdote (2011), "Did the stimulus stimulate? real time estimates of the effects of the american recovery and reinvestment act." Working Paper 16759, National Bureau of Economic Research.
- Foster, Lucia, John Haltiwanger, and Cornell John Krizan (2001), "Aggregate productivity growth: Lessons from microeconomic evidence." In *New developments in productivity analysis*, 303–372, University of Chicago Press.
- Gabler, Alain and Markus Poschke (2013), "Experimentation by firms, distortions, and aggregate productivity." *Review of Economic Dynamics*, 16, 26–38.
- Gawande, Kishore and Usree Bandyopadhyay (2000), "Is protection for sale? evidence on the grossman-helpman theory of endogenous protection." *The Review of Economics and Statistics*, 82, 139–152.
- Gottlieb, Joshua D, Richard R Townsend, and Ting Xu (2016), "Experimenting with entrepreneurship: The effect of job-protected leave." NBER Working Paper #22446.
- Guzman, Jorge and Scott Stern (2016), "Nowcasting and placecasting entrepreneurial quality and performance." In *Measuring Entrepreneurial Businesses:* Current Knowledge and Challenges, University of Chicago Press.
- Haltiwanger, John, Henry R Hyatt, Erika McEntarfer, Liliana D Sousa, and Stephen Tibbets (2014), "Firm age and size in the longitudinal employer-household dynamics data." U.S. Census Bureau Center for Economic Studies Working Paper #CES-14-16.
- Haltiwanger, John, Ron Jarmin, Robert Kulick, and Javier Miranda (2016), "High growth young firms: Contribution to job, output, and productivity growth." In *Measuring Entrepreneurial Businesses: Current Knowledge and Challenges*, University of Chicago Press.

- Haltiwanger, John, Ron Jarmin, and Javier Miranda (2013), "Who creates jobs? Small versus large versus young." *Review of Economics and Statistics*, 95, 347–361.
- Hamilton, Barton H (2000), "Does entrepreneurship pay? An empirical analysis of the returns to self-employment." *Journal of Political Economy*, 108, 604–631.
- Hendricks, Kevin B and Vinod R Singhal (2005), "Association between supply chain glitches and operating performance." *Management Science*, 51, 695.
- Holmes, Thomas J, David K Levine, and James A Schmitz Jr (2012), "Monopoly and the incentive to innovate when adoption involves switchover disruptions." *American Economic Journal: Microeconomics*, 4, 1–33.
- Holtz-Eakin, Douglas, David Joulfaian, and Harvey S. Rosen (1994), "Sticking it out: Entrepreneurial survival and liquidity constraints." *Journal of Political economy*, 102, 53–75.
- Hombert, Johan, Antoinette Schoar, David Alexandre Sraer, and David Thesmar (2017), "Can unemployment insurance spur entrepreneurial activity? Evidence from france." Working Paper, HEC Paris, MIT and UC Berkeley.
- Hsieh, Chang-Tai and Peter J. Klenow (2014), "The life cycle of plants in india and mexico." The Quarterly Journal of Economics, 129, 1035–1084.
- Hurst, Erik and Benjamin W. Pugsley (2011), "What do small businesses do?" Brookings Papers on Economic Activity, 2011, 73–118.
- Hurst, Erik G and Benjamin W. Pugsley (2016), "Wealth, tastes, and entrepreneurial choice." In *Measuring Entrepreneurial Businesses: Current Knowledge and Challenges*, University of Chicago Press.
- Imbens, Guido and Donald Rubin (2015), Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction, 1 edition. Cambridge University Press.
- Imrohoroglu, Ayse and Selale Tuzel (2014), "Firm-level productivity, risk, and return." *Management Science*, 60, 2073–2090.
- Jarmin, Ron and Javier Miranda (2002), "The longitudinal business database." CES working paper 02-17.
- Jovanovic, Boyan (1982), "Selection and the evolution of industry." *Econometrica*, 50, 649–70.
- Kaiser, Ulrich and Nikolaj Malchow-Møller (2011), "Is self-employment really a bad experience?: The effects of previous self-employment on subsequent wage-employment wages." *Journal of Business Venturing*, 26, 572–588.
- Kang, Karam (2015), "Policy influence and private returns from lobbying in the energy sector." Review of Economic Studies, 83, 269–305.

- Kerr, Sari Pekkala and William R. Kerr (2016), "Immigrant entrepreneurship." In *Measuring Entrepreneurial Businesses: Current Knowledge and Challenges*, University of Chicago Press.
- Kerr, William R, William F Lincoln, and Prachi Mishra (2014a), "The dynamics of firm lobbying." *American Economic Journal: Economic Policy*, 6, 343–79.
- Kerr, William R., Ramana Nanda, and Matthew Rhodes-Kropf (2014b), "Entrepreneurship as experimentation." *The Journal of Economic Perspectives*, 28, 25–48.
- King, Robert G and Ross Levine (1993), "Finance, entrepreneurship and growth." Journal of Monetary economics, 32, 513–542.
- Laasch, Oliver and Roger Conaway (2009), "Gender differences in preferences." Journal of Economic Literature, 47, 448–474.
- Levinsohn, James and Amil Petrin (2003), "Estimating production functions using inputs to control for unobservables." Review of Economic Studies, 70, 317–342.
- Liebman, Jeffrey and Neale Mahoney (2013), "Do expiring budgets lead to wasteful year-end spending? evidence from federal procurement." Working Paper 19481, National Bureau of Economic Research.
- Lucas, Robert E. (1978), "On the size distribution of business firms." *Bell Journal of Economics*, 9, 508–523.
- Luttmer, Erzo G.J. (2012), "Technology diffusion and growth." *Journal of Economic Theory*, 147, 602–622.
- Manso, Gustavo (2016), "Experimentation and the returns to entrepreneurship." The Review of Financial Studies, 29, 2319–2340.
- Meade, Janet and Shihong Li (2015), "Strategic corporate tax lobbying." The Journal of the American Taxation Association, 37, 23–48.
- Midrigan, Virgiliu and Daniel Yi Xu (2014), "Finance and misallocation: Evidence from plant-level data." *American Economic Review*, 104, 422–58.
- Moll, Benjamin (2014), "Productivity losses from financial frictions: can self-financing undo capital misallocation?" *American Economic Review*, 104, 3186–3221.
- Moreira, Sara (2017), "Firm dynamics, persistent effects of entry conditions, and business cycles." U.S. Census Bureau Center for Economic Studies Working Paper #CES-17-29.
- Polkovnichenko, Valery (2003), "Human capital and the private equity premium." Review of Economic Dynamics, 6, 831–845.

- Pugsley, Benjamin W and Aysegul Sahin (2015), "Grown-up business cycles." U.S. Census Bureau Center for Economic Studies Working Paper #CES-15-33.
- Richter, Brian, Krislet Samphantharak, and Jeffrey Timmons (2009), "Lobbying and taxes." American Journal of Political Science, 53, 893–909.
- Schoar, Antoinette (2010), "The divide between subsistence and transformational entrepreneurship." In *Innovation Policy and the Economy*, volume 10, 57–81, University of Chicago Press.
- Schumpeter, Joseph A (1942), Capitalism, socialism and democracy. Harper and Brothers.
- Sedlácek, Petr and Vincent Sterk (2017), "The growth potential of startups over the business cycle." *American Economic Review*, 107, 3182–3210.
- Törnqvist, Leo, Pentti Vartia, and Yrjö O. Vartia (1985), "How should relative changes be measured?" *The American Statistician*, 39, 43–46.
- Vereshchagina, Galina and Hugo A Hopenhayn (2009), "Risk taking by entrepreneurs." American Economic Review, 99, 1808–30.
- Warren, Patrick L (2014), "Contracting officer workload, incomplete contracting, and contractual terms." The RAND Journal of Economics, 45, 395–421.
- Williams, Donald R. (2000), "Consequences of self-employment for women and men in the united states." *Labour Economics*, 7, 665–687.
- Wilson, Daniel J. (2012), "Fiscal spending jobs multiplier: Evidence from the 2009 american recovery and reinvestment act." *American Economic Journal: Economic Policy*, 4, 251–282.
- Wooldridge, Jeffrey (2009), "On estimating firm-level production functions using proxy variables for unobservables." *Economic Letters*, 104, 112–114.