The Promise and Peril of Entrepreneurship: Job Creation and Survival among U.S. Startups (longer version published as *MIT Press* 2023 book)

Robert Fairlie - UCLA and NBER
rfairlie@ucsc.edu

Zachary Kroff - U.S. Census Bureau
zachary.kroff@census.gov

Javier Miranda – Halle Institute for Economic Research (IWH) and Friedrich Schiller University
Javier.Miranda@iwh-halle.de

Nikolas Zolas - U.S. Census Bureau
nikolas.j.zolas@census.gov

March 2023

Abstract: Entrepreneurship is promoted around the world by governments and policymakers, but surprisingly, we do not know the answers to three fundamental questions important for basic economic welfare and policy calculations: i) How many jobs does an entrepreneur create?, ii) Do these jobs disappear quickly?, and iii) How many entrepreneurial firms survive each year after startup? We provide some of the first answers to these questions by creating a new panel dataset that captures the universe of U.S. startups, the Comprehensive Startup Panel (CSP). Using the broadest definition possible, we find that the average entrepreneur creates 0.74 jobs at the first year after startup and continues to employ 0.63 workers five years after startup (using the broadest definition possible). In total, the average annual cohort of 4.1 million startups in the United States creates 3.0 million jobs in the first year after startup. These jobs are important because net job creation by all other businesses is negative on average. We also find low survival rates – only 47 percent of startups survive two years and only 33 percent survive five years, with no industries being immune. We also find that non-employer startups make sizeable contributions to employment several years after startup (one seventh of jobs seven years later). Using a more restrictive definition of startups (which excludes sole proprietors without EIN startups) we find an upper-bound job creation rate of 2.6 jobs per startup after one year and a survival rate of 45 percent five years later. Even with this more restrictive definition we find job creation per startup and survival rates that are notably lower and less optimistic than the oft-cited numbers released by the federal government that focus exclusively on new employer businesses.

Keywords: entrepreneurship, startups, job creation, survival rates

Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed. We thank Ronnie Chatterji, Brain Headd, Ross Levine, Ben Pugsley, David Robinson, Bulent Unel, and seminar participants at Duke University, Stanford University, the National Academies of Science, Census Bureau, Small Business Administration, Louisiana State University, University of Oklahoma, University of Nevada, Reno, the FSRDC Conference at UCLA, and the NBER Entrepreneurship workshop for helpful comments and suggestions.

I. Introduction

Federal, state and local governments spend billions of dollars each year on incubators, training programs, loan programs, tax breaks, and investor incentives to encourage business creation with one of the primary goals being to create jobs.[1] Expenditures are generally made, however, without measurement of the benefits in terms of the number of businesses, jobs or total payroll created relative to the costs of these programs.[2] Surprisingly, relevant to any benefit to cost calculation, we do not know the answers to the fundamental questions of how many jobs does the average entrepreneur create or what percentage of entrepreneurial firms survive one, two, or more years later?[3] But, in spite of the perception of high failure rates and uncertainty over job creation, the popularity of public policies to foster entrepreneurs continues to grow around the world (OECD 2017).

Measuring job creation and survival among entrepreneurs is also important for better understanding the fundamental nature of entrepreneurship. Is entrepreneurship about creating innovative products, services and jobs (i.e. Schumpeterian entrepreneurship), or is it becoming increasingly about job independence (i.e. being one's own boss), contract/consulting work, schedule flexibility, or being part of the gig economy (Hamilton 2000; Hurst and Pugsley 2011; Levine and Rubinstein 2016; Katz and Krueger 2017)? In other words, is entrepreneurship about creating *jobs* or creating *a job*? Is it about creating a long-term job or just a temporary job? And,

---

[1] The U.S. Small Business Administration, for example, administers several programs to support small businesses, including loan guaranty, training, federal contracting, and other programs.
[2] The evidence on the effectiveness of programs to spur entrepreneurship is mixed. For a few recent examples, see Brown and Earle (2016) on SBA lending programs, Chrisman (2016) on Small Business Development Centers (SBDCs), Lerner (1999, 2009) on the Small Business Innovation Research (SBIR) and other public venturing programs, Fairlie, Karlan and Zinman (2015) on entrepreneurship training, Huber et al. (2014) on entrepreneurship education, Chatterji, Chay and Fairlie (2014) on contracting programs, and Mathematica (2017) on SEA programs.
[3] Recently, the U.S. Government Accounting Office (GAO) recommended that the SBA use job creation and other outcome-based measures to evaluate the effectiveness of its programs (CRS 2017). Also, SBDCs, which are partly funded by SBA, measure job creation resulting from their training and assistance centers around the country from surveys of participants (Chrisman 2016).

finally, is it about taking risks (Knight 1921; Kihlstrom and Laffont 1979) or diversifying risk (i.e. through multiple job activities)?

In this paper, we create and use a new compilation of administrative data on the universe of business startups in the United States to answer three fundamental questions about entrepreneurship, job creation and survival.[4] Using the newly created comprehensive startup panel (CSP) dataset, we provide the some of the first answers to: i) How many jobs are created by the average entrepreneur?, ii) Do these jobs last over time?, and iii) How many entrepreneurial firms survive each year after startup?[5] In addition to providing answers to these fundamental questions about entrepreneurship, we test several proposed explanations for the underlying causes of patterns in entrepreneurial job creation and survival.

To track job creation and survival over time by every business startup in the U.S. economy, we link the universe of non-employer firms to the universe of employer firms in the Longitudinal Business Database (LBD). To create the necessary links between non-employer and employer definitions, we update and expand the Integrated Longitudinal Business Database (iLBD). In the process, we create a new compilation of existing administrative data that covers the universe of startups over several follow-up years.

The federal government releases two oft-cited and influential statistics on startup job creation and survival rates – 6 new jobs per startup and 50 percent survival after 5 years (U.S. Small Business Administration 2017), however, the underlying data used to create these numbers include only new *employer* businesses or establishments and do not look back for previous

[4] Fairlie et al. (2019) also describe the new dataset that we created, and provide a couple of illustrative examples of what can be done with the data using the first "beta" cohort from 1997.
[5] There is no universally agreed upon or official definition of "entrepreneurs" in government data or research (Decker et al. 2015; Parker 2009). The U.S. Bureau of Labor Statistics (BLS) publishes statistics on "entrepreneurs" that are defined as new employer establishments. We follow this approach of defining entrepreneurship as new businesses or business startups.

existence as a non-employer startup.[6] In this paper, we measure entrepreneurial job creation and survival rates starting with *every* startup in the United States and then consider more restrictive definitions. The classic studies of entrepreneurship, such as Knight (1921), Schumpeter (1934) and more recently Kihlstrom and Laffont (1979), Jovanovic (1982) and Evans and Jovanovic (1989), that have helped establish entrepreneurship as an independent field in economics do not limit the definition of entrepreneurship by employer, incorporation or any other status. This is important because estimates of job creation per *employer* startup will greatly overestimate job creation per *any* startup. Non-employer firms comprise the majority of both startups and total firms in the United States. Also, employer firms will be misclassified as new businesses if they actually started several years earlier with no employees, and jobs eventually created by non-employer startups will be credited to the wrong startup cohort.[7] The inclusion of non-employer startups is also important for correctly measuring survival rates over time because many non-employer firms exit before ever hiring an employee. In general, ignoring the non-employer history of firms is likely to miss important early entrepreneurial dynamics that are crucial to understanding the relationships between entrepreneurship, job creation, and survival.

The analysis of the new administrative dataset on the universe of U.S. startups (CSP) provides several findings that shed light on fundamental, and surprisingly unanswered, questions regarding the job creation potential and survival of entrepreneurs. Using the broadest definition possible, we find that the average entrepreneur creates 0.74 jobs at the first year after startup, and

[6] The previous research on job creation among businesses focuses almost exclusively on employer firms (e.g. Haltiwanger, Jarmin and Miranda 2013; Kulick, Haltiwanger, Jarmin; Miranda 2016; Glaeser, Kerr and Ponzetto 2010; Glaeser, Kerr and Kerr 2015; Tracy 2011; Decker et al. 2014; Garcia-Macia, Hsieh, and Klenow 2016). Data on non-employer startups are more difficult to find and are not reported by age of business over time. Among OECD countries, the United States is one of the few countries that does not report non-employer business creation rates (see OECD 2017, Figure 4.1 for example).

[7] A simple example illustrates this point. If a business with no employees starts in 2001 and hires its first employee three years later, data focusing on employer firms will capture this business as a startup in 2004. The jobs created by this business will also be attributed to startup cohort 2004 instead of startup cohort 2001.

employs 0.63 workers five years later and 0.57 workers seven years later.[8] The average annual cohort of 4.1 million startups in the United States creates a total of 3.0 million jobs in the first year after startup and employs 2.6 million workers five years later. Without these jobs created by startups, net job creation would be negative and total nearly 2 million job losses per year. We also find that survival rates are extremely low among the universe of startups with a large shake out occurring in the years immediately after startup. After one year, only 59 percent of startups survive and after two years only 47 percent survive. The decline in survival rates starts to taper off, and after five years 33 percent of startups survive. Surprisingly, survival rates are remarkably consistent across industries, differing from job creation, which varies substantially across industries.

We also test several competing hypotheses regarding the underlying causes of job creation dynamics. Exploring the rich heterogeneity in job creation dynamics across the universe of startups, we find that rapid job growth over time among surviving startups mostly offsets the negative influence from extremely high exit rates. Surviving startups hire an average of 1.9 employees five years after startup and 2.0 employees seven years after startup (compared with hiring 1.3 employees in the first year after startup).

When we turn to investigating heterogeneity based on business type at startup we find that non-employer startups make substantial contributions to job creation: an average of nearly 320,000 jobs seven years after startup representing one-seventh of the total 2.3 million jobs created by all startups. We also examine whether the inclusive definition of entrepreneurship explains why job creation is low and survival rates are low. Because of concerns over including

[8] We focus here on net employment levels over time per startup. The data do not provide information on gross flows during the year or whether employees were previous unemployed, out of the labor force, or working for another business.

all non-employer business entities, we also experiment with more restrictive definitions. Using a more restrictive definition of entrepreneurship that includes employer, incorporated, partnership, S-corporation startups and sole proprietorships with EINs, but excludes sole proprietorships without EINs, we find that the average entrepreneur creates 2.56 jobs in the first year after startup and 2.15 jobs five years later. Although job creation is higher among this select group of startups, survival rates remain low, with 64 percent surviving after 2 years and 45 percent surviving after 5 years. We view these levels of job creation per entrepreneur and survival rates as upper bounds. These findings paint a picture of the U.S. entrepreneur as one who creates few jobs and experiences high exit rates, but when surviving grows steadily.

Our paper is related to several influential studies on job creation among small businesses. These studies focus on identifying the share of jobs created by small or young businesses relative to large or older businesses. Starting with the seminal study by Birch (1979) showing that small businesses are the principal driver of job creation in the U.S. economy, there has been considerable interest in job creation among entrepreneurs. Recent evidence indicates that young and high-impact businesses (defined as having high rates of growth in sales and employment) disproportionally contribute jobs in the economy (Haltiwanger, Jarmin and Miranda 2013; Kulick, Haltiwanger, Jarmin, Miranda 2016; Tracy 2011). A few recent studies also examine the relationship and growth patterns between non-employer to employer businesses. These studies find, for example, that non-employers have startup rates that are nearly three times the startup rates of employer firms, a significant number of new employer firms start as non-employer firms, and if non-employer startups hire, the bulk of hiring occurs in the first few years of existence

(Davis et al. 2007; Acs, Headd and Agwara 2009; Fairlie and Miranda 2017).[9] Another strand of literature examines entrepreneurial survival rates (e.g. Phillips and Kirchhoff 1989; Audretsch 1991; Audretsch Mahmood 1995; Robb and Farhat 2013; BLS 2017; SBA 2017). Recent estimates from the BLS (2017) indicate that 48 percent of new employer establishments survive after five years. None of these previous studies, however, use the full universe of startups and the entire history of non-employer and employer status to measure entrepreneurial job creation and survival.

The absence of non-employer startups represents an important omission in the literature because of the nearly 3.7 million non-employer business entities created each year in the United States and that create one-seventh of all jobs among startups seven years later. Furthermore, we find that in the average cohort of firms hiring their first employee that roughly one-fifth had a non-employer history. Most importantly, however, we find job creation rates per startup and survival rates that are substantially lower and less optimistic than the numbers presented by the federal government that focus on employer startups.

Our paper also contributes more generally to the rapidly growing literature on entrepreneurship. This emerging field has struggled with finding adequate data and agreement on definitions of entrepreneurship. We create a new compilation of administrative data that captures the universe of business startups in the United States. The new U.S. startup panel dataset (CSP) allows for the exploration of many questions around entrepreneurial job creation, survival and growth, and allow for substantial flexibility in how one defines entrepreneurship. Additionally, the massive number of observations – an average of 4.1 million startups *each* year and data

[9] Using income tax return data, Carroll et al. (2000) find that among their sample of individuals who were sole proprietors in both 1985 and 1988 roughly one-third hired workers. They also found that 9 percent took on workers and 22 percent stopped hiring workers between the two years.

covering 1995 to 2018 – allows for unprecedented detail in analyzing different types of entrepreneurs. The administrative panel data also eliminate concerns over survivor, recall and attrition bias. The new compilation of administrative data on the universe of startups created, described and examined here may be useful for future research.

The rest of the paper proceeds as follows. In section II, we describe the data on startups. Section III presents the main empirical results on job creation, and Section IV presents the main empirical results on survival. Section V explores several potential explanations for underlying job creation and survival dynamics. Section VI explores questions related to heterogeneity in startup types and an alternative definition of entrepreneurship. Section VII concludes.

II. Data on Startups

To measure entrepreneurial job creation and survival rates, we create and analyze a new compilation of administrative panel data on the universe of business startups in the United States. We link the universe of non-employer firms to the universe of employer firms in the Longitudinal Business Database (LBD). To create the necessary links between non-employer and employer definitions, we update and expand the Integrated Longitudinal Business Database (iLBD). Beta versions of the iLBD have been used in previous work (i.e. Davis et al 2007; Fairlie and Miranda 2017) to study transitions between non-employer and employer business units, but these analyses were only for one or two years, a subset of industries, and/or not focused on startups. A beta version of the iLBD was first introduced by Davis et al (2007) to study transitions between non-employer and employer business units for a subset of industries in 1992 and 1994. Subsequently, Fairlie and Miranda (2017) created another beta version with the 1997 non-employer startup cohort to examine transitions to employer businesses. And, most recently, Fairlie et al. (2019) also provide a description of the new dataset for a special issue on entrepreneurship in the *ILR Review* Using only the first 1997 "beta" cohort, the article provides a few illustrative examples of what can be done with the data in future research, but the linkages for that first cohort have been revised and the new numbers are showing fewer startups.

For this paper, we combine the underlying confidential and restricted access iLBD and LBD microdata for several startup cohorts (i.e. 1995 to 2011 startup cohorts). The new startup panel dataset created here follows each startup cohort for sevenl years after startup covering the years 1996 to 2018. The underlying iLBD is sourced from administrative income and payroll filings. The iLBD provides information on the universe of non-employer firms as well as a set of identifiers that allow connecting them to the universe of employer firms in the LBD.[10] The combination of the iLBD and the LBD then allows us to explore the connections between the two universes including transitions between the two. The resulting startup panel dataset provides annual information that allows all businesses, employer and non-employer, to be followed over time. Because the underlying data contain links between the employer and non-employer universes, it is possible to accurately identify the point of startup for business units and each annual cohort of startups can be followed over time. A business entity can only be included in the startup cohort for that year if it is not found in previous years of the non-employer or employer universes. Additionally, mergers and acquisitions of existing firms are not identified as startups, but are instead treated as a continuation of the initial business.

The new U.S. startup panel dataset provides several advantages for studying the universe of startups. First, the non-employers that hire employees in some future year are included and not just firms that start with employees. Second, the data provide information at the point in time of

[10] For details about the LBD see Jarmin and Miranda (2002).

hiring instead of less reliable retrospective information that might be collected through surveys.[11] Another advantage of the startup panel dataset relative to many other datasets is that it contains administrative data on the universe of business units, and thus suffers little from attrition problems. Examining job creation per startup over time using cross-sectional data is problematic because only surviving businesses are included. Attrition with longitudinal survey data is also a serious concern when exploring entrepreneurial dynamics. Even low annual rates of attrition lead to high rates of long-term attrition. It is thus difficult to accurately estimate job creation and exit rates over time by a cohort of startups using survey data.

A careful examination of job creation and survival among the universe of startups requires identifying transitions between non-employer and employer status.[12] The employer and non-employer administrative data are kept separate at the Census Bureau making it difficult to identify these transitions. Businesses in the United States are required to file separately income and payroll (employment) taxes. However, businesses may or may not use the same tax identifiers when filing their income and payroll reports.[13] Since the transition to employer status can only be identified by linking the income filing to the payroll filing, this can lead to broken linkages between the two. To resolve this problem the iLBD links the employer and non-employer businesses units by a variety of identifiers including the EIN, the Protected Identification Key (PIK), and the name and address of the owner or business.[14] These enhanced linkages are tracked in the iLBD through longitudinal business identifiers, iLBDNUMs.

[11] Employment in tax filings covers the payroll period of March 12th.

[12] The Kauffman Firm Survey (KFS) which provides data on a sample of roughly 4,000 non-employer and employer startups has also been used to study non-employer to employer transitions (Fairlie and Miranda 2016). Because of the underlying sampling frame of the KFS, however, the sample is skewed towards including employer firms.

[13] Non-employer businesses do not file payroll taxes.

[14] The PIK is an individual identifier that replaces the Social Security Number (SSN) in all files at the U.S. Census Bureau. The PIK ensures no Census employee or researcher has access to SSNs. The linkage between the non-employer and the employer universes makes use of the name of the business and the tax identifiers; the EIN and the PIK. If any of these change then it might not be possible to form a link. This is more likely when there is a change in

The U.S. startup panel dataset includes every non-employer business that files taxes or a tax report. One concern is that the startup panel dataset contains a large number of business activities that have no intention of hiring employees and represent consulting, contracting or hobby activities. But, these data provide a useful view of the universe of non-employer business units, and we are able to identify more growth-oriented businesses by conditioning on a few administrative variables. The data contain information on the legal form; sole proprietor, partnership or corporation, the type of tax identifier; EIN or PIK, their revenue, and industry of activity in addition to information on startup year, employment, payroll, exits, and geographical location. We explore this possibility and discuss the issues in detail below by examining populations with and without business entities registered as sole proprietors with no employees at startup.

The new U.S. startup panel dataset, underlying iLBD and LBD data, and compilation code are confidential and restricted-access. They are accessible to researchers with an approval through Federal Statistical Research Data Centers (FSRDC).[15]

III. Job Creation by Startups

On average, 4.1 million businesses are started each year in the United States. The 4.1 million startups create an average of 3.0 million jobs in the first startup year. Figure 1 displays total job creation for the entire startup cohort following job creation from the first year after startup to seven follow-up years after startup (see also Table 1). The total number of jobs created

legal form of organization at the time of the transition. For additional discussion of these issues see Davis et al. (2007).

[15] Information about the application process and more generally about FSRDCs is available at http://www.census.gov/fsrdc. Also, see Fairlie et al. (2019) for more information.

(net) declines to 2.6 million by the fifth follow-up year and 2.3 million by the seventh follow-up year.

The 3.0 million jobs created by startups in the first year after startup represents roughly 3 percent of total employment by all businesses. Figure 2 displays the share of total U.S. employment created by startups in each year since startup. To calculate the startup employment share, total employment is averaged across all of the years following each cohort. Total U.S. employment ranged from 112 million to 118 million over the corresponding years.[16] In the first follow-up year job creation among startups represents 2.7 percent of total employment by all firms. After five years the share of total U.S. employment by the average startup cohort is 2.2 percent, declining to 2.0 percent after seven years.

Although startups represent a small share of total employment they represent a large percentage of net job creation each year, especially in the first two years of existence. Figure 3 displays net job creation by startups and all other firms over time. Net job creation for all other firms is negative in every year, averaging 1.9 million job losses per year. Without the job creation contributions of startups through their first seven years of existence, net job creation would be negative in the United States.

These statistics indicate the importance of young firms to job creation in the United States. Using LBD microdata and focusing on the 2005 cohort, Haltiwanger, Jarmin and Miranda (2013) find that new (e.g. age 0) employer firms created 3.5 million net new jobs compared with a loss of 1 million net jobs for firms of all other ages (i.e. age 1 and over) combined. These findings are similar to our estimates for the initial year contribution to job creation for the average startup cohort (which includes all startups).

[16] Total employment by all U.S. businesses is from U.S. Census Bureau (2020).

Job Creation per Startup

We turn to examining the questions of i) How many jobs are created by each entrepreneur?, and ii) Do these jobs last over time? The answers to these questions are fundamental to how we think about entrepreneurship from policy, economic welfare, and theoretical perspectives. Figure 4 displays the number of jobs created by the average startup over seven follow-up years (see also Table 1). The average entrepreneur creates 0.74 jobs in the first year after startup. Average employment per startup does not disappear quickly: it is 0.63 five years after startup and 0.57 seven years after startup. All startups in a cohort are included in the denominator even if they exit. Administrative panel data on the universe of startups is essential for this calculation because we do not have slippage in job creation through survey response attrition or survival bias.

These patterns of job creation by the average startup reveal two important findings. First, job creation levels per entrepreneur are relatively low. Startups create an average of roughly 3/4 job each. Second, although total levels of job creation might be viewed as relatively low, the jobs created by entrepreneurs do not disappear right away. The number of jobs created by the average startup only declines by 0.11 jobs, from 0.74 jobs to 0.63 jobs 5 years later. Net job creation by startups is relatively long lasting. These levels of job creation are considerably lower than the 6 jobs per employer startup oft-cited in federal statistics (e.g. SBA 2017). We discuss this comparison in more detail below.

Payroll per Employee

From administrative payroll records, we can also examine how much employees are paid on average at each startup. Figure 4.2 displays average payroll per employee by years since startup. All payroll statistics are adjusted for inflation and reported in 2015 dollars. The average earnings per employee is $34,700 in the first year after startup. Earnings per employee steadily increases in each follow-up year. By the fifth follow-up year, earnings per employee are $38,600, rising to $40,000 by year 7. Average salaries at startups are lower than more established businesses. Establishment level data from the U.S. Census Bureau indicate that over a similar time period, all business establishments in the United States paid an average of roughly $45,000 per employee. The finding of lower wages paid at young firms relative to older firms is in line with previous work focusing on employer businesses (Brown and Medoff, 2003; Haltiwanger et al., 2012).

Regression Results

We now turn to an exploration of the factors associated with job creation among the universe of startups. To investigate this question we estimate the following equation:

(1) $Y_{it} = \alpha + \sum_{s=2}^{7} \delta^s D_i^s + \sum_{k=1}^{19} \beta^k I_i^k + \gamma_t + u_i + \epsilon_{it}$

where Y_{it} is employment at firm i in year t, D_i^s is a set of dummy variables for years since startup for firm *i*, the excluded or reference year is the first year after startup), I_i^k is a set of dummy variables for industries, γ_t are fixed effects capturing each annual startup cohort, , and $u_i + \epsilon_{it}$ is a composite error term. The main parameters of interest are δ^s which captures job creation in follow-up year *s* relative to the omitted first follow-up year and β^k which captures job creation for industry *k*. Standard errors are clustered at the firm level to account for multiple observations per startup.

To estimate Equation (1), we include the full sample of firm/year observations even after exits. In the case of exits, we set employment to zero and include observations for all subsequent years for that firm. We thus treat a non-surviving startup as having no employment, which is similar to the idea of net job creation per startup over time. Because we have administrative data on all startups over time, we can rule out the concern that sample attrition is masking exits.

Although not reported in the table, we also control for the differences in job creation across startup cohorts. We do not find major differences across startup cohorts, and do not focus on that here.

Specification 1 of Table 2 reports estimates of Equation (1). The sample size for this regression is more than 550 million observations and the unit of analysis is the at the business-year level. Even clustering at the business level, the standard errors on all reported coefficient estimates are extremely small and all coefficients are statistically significant at the 0.01 level (and even the 0.001 level).

After controlling for startup cohort years and industries, we find that the average number of employees per startup decreases over time. The second follow-up year after startup job creation is slightly lower than first year after startup (-0.019 jobs). For reference, the first year after startup has a job creation level of 0.74. The fifth year after startup shows a decline of 0.11 jobs and the seventh year shows a decline of 0.17 jobs relative to the first followup year. These estimates are similar to those implied by the patterns displayed in Figure 4. The relatively constant or slight declining pattern across years since startup is not driven by differences in industries or startup cohorts.

Table 2 also reports coefficients by industry. Job creation levels differ substantially across industries (unadjusted job creation levels by industry are reported in Appendix Table 1).

In all regressions reported here and below, “Other Services” is the left out category.. A few industries have employment levels that are higher: Management (+2.3), Manufacturing (+2.9), and Accommodation and Food Services (+3.2). Industries with the lowest job creation levels in addition to Other Services, which is the third lowest level, are Agriculture (-0.04), Real Estate (-0.05), Arts, Entertainment and Recreation (0.06), Transportation (0.13), and Professional (0.14).

IV. Survival Rates

Using the universe of startups, we explore the third fundamental question: How many entrepreneurial firms survive each year after inception? Figure 5 reports the number of startups surviving over the seven-year follow-up period from the representative startup cohort (see also Table 1). In all of our analyses of survival rates, we do not distinguish between business closures and failures.[17] Because of data limitations and to follow federal government convention, we focus on survival of the business entity. Of the 4.1 million startups, a large number disappear in each follow-up year. Roughly 40% of startups exit by year 1 (2.4 million startups survive) and 53% exit by year 2 (1.9 million startups survive). After this initial shakeout, the exit rate starts to slow down. After five years, 1.4 million startups remain in operation and after seven years, 1.2 million startups remain in operation. Figure 5.2 displays essentially the same information as Figure 5, but as survival rates for the average startup cohort. These survival rates are substantially lower than the 50 percent after 5 years number referenced widely in policy, academic, and popular circles (e.g. SBA 2017). We discuss the comparison further in the next section.

[17] See Headd (2003) and Parker (2009) for discussions on the limitations of focusing solely on business exits.

Regressions for Survival Rates

We now turn to an exploration of the factors associated with survival among the universe of startups. To investigate this question we estimate Equation (1) where Y_{it} is now survival of firm i in year t (equal to 0 or 1). As before, the main parameters of interest are δ^s and β^k. Note that survival, $Y_{it} = 1$, in the startup year for all businesses, and the startup year is the excluded or reference category in the regression. Standard errors are clustered at the firm level to account for multiple observations per startup. Regressions are estimated with a linear probability model (marginal effects estimates for logit and probit models are similar and not reported).

To estimate Equation (1) we include the full sample of firm/year observations even after exits. In the case of exits, we set survival equal to zero and include observations for all subsequent years for that firm. Alternatively, we could estimate a hazard model for the length of the survival spell, but the typical advantages of hazard models such as addressing left and right censoring of spells and having multiple spells do not apply here. To simplify estimation with the restricted and confidential data with a large sample we estimate the regressions with a linear probability model (OLS) using all possible observations over the 8-year window for each startup. Similar to the employment regressions, the administrative data on startups rule out the concern that sample attrition is masking exits.

Specification 1 of Table 3 reports estimates of Equation (2). The sample size for this regression is roughly 550 million observations and the unit of analysis is at the business-year level. Even clustering at the business level, the standard errors on all reported coefficient estimates are extremely small and all coefficients are statistically significant at the 0.01 level (and even the 0.001 level).

After controlling for startup years and industries, we find that survival rates decrease sharply in the first couple of years and then taper off. Survival rates drop by 0.42 in the first follow-up year after startup. Two years after startup survival rates have dropped by 0.54 from the startup year. After five years post startup the survival rate drops 0.68 from the startup year. Controlling for startup year and industry differences does not change the finding of a strong downward trend in survival rates over years since startup.

In contrast to job creation rates by industry, survival rates do not differ substantially by industry (unadjusted survival rates by industry are reported in Appendix Table 2). Other Services is the omitted category. Most industries do not differ substantially in either direction from this omitted category. The main exceptions are that the survival rate is 4.1 percentage points lower for Information, 4.5 percentage points lower for Administrative and Support, and 12.4 percentage points higher for Real Estate. But, overall, the differences are not large and survival probabilities are largely independent of industry.

V. Heterogeneity in the Underlying Dynamics of Job Creation by Startups

In this section, we explore several questions related to the underlying dynamics of job creation and survival among startups. First, we explore the question of why does the average number of jobs per entrepreneur remain relatively constant over years since startup. Patterns in the average number of jobs created per entrepreneur conceal a substantial amount of heterogeneity in the underlying dynamics of job creation by startups, which might hold the answer. We explore four potential explanations for the relatively flat pattern: i) a large percentage of startups survive and these startups grow slowly, ii) a large percentage of startups survive and a small percentage of the survivors grow rapidly, iii) a small percentage of startups survive and a large percent of survivors grow steadily, and iv) a small percentage of startups survive and a small percentage of survivors grow very rapidly.

Because such a large percentage of startups exit in the first few years of operation, we can likely rule out the first two explanations due to the relatively flat level of job creation over time. To examine this directly, however, Figure 6 displays job creation per surviving startup and job creation per all startups over time. Job creation levels per *surviving* startup grow rapidly over time. One year after startup, the average startup employs 0.74 workers. Among surviving startups one year later, the average startup hires 1.27 employee. Seven years later, surviving startups hire an average of 1.96 employees each. These results are consistent with the strong "up or out" pattern of startup dynamics first noted among employer startups in Haltiwanger, Jarmin and Miranda (2013).

Overall, the strongly opposing patterns of increasing job creation among surviving startups and decreasing survival rates among startups mostly cancel each other out. This explains why total job creation from the first year after startup to five years after startup only changes by 0.11 employees (15 percent). To the seven years after startup it changes by 0.17 employees (23 percent).

Regressions Conditioning on Survival

Given the strong pattern of employment growth among surviving startups, we examine the determinants of job creation conditional on survival. Are the characteristics associated with job creation per startup different when focusing on only *surviving* startups? To investigate this question we estimate Equation (1) including only surviving startups. We condition on survival in that year. In this case, we exclude all annual observations after firm exit. We are thus estimating a model that measures net job creation per *surviving* startup.[18]

Specification 2 of Table 2 reports estimates for the conditional model (i.e. includes only surviving firm-years in the sample). The same sets of variables and controls are included in the regression. The sample size is now much smaller do to exits. Similar to the unconditional sample estimates we do not find a strong pattern across startup cohorts. Regression estimates from the survivor sample of startups indicate a different time-series pattern for job creation than for the unconditional model which includes all observations. Conditioning on survival, the number of jobs per startup increases steadily from the first full year of the business. Two years after startup, average employment increases by 0.30 compared to the first year after startup. Five years after startup, average employment increases by 0.69 and seven years after startup employment increases by 0.77. This is due to the removal of startups that exited by the time of that follow-up year and thus are not included in the sample (or denominator) anymore. This also is the rate after we control for other potentially confounding factors, such as cohort year and sector.

The patterns across major industry groups generally follow the same patterns as the unconditional sample estimates. We continue to find that the sectors with the highest levels of conditional job creation are Accommodation and Food Services (+6.33), Manufacturing (+5.74), and Management (+4.69). This should be interpreted as an overall level shift of the growth pattern summarized in the regression. Sectors with the lowest levels of job creation conditional on survival in addition to Other Services, which is the third-lowest level, are Agriculture and

[18] We considered estimating a selection model that simultaneously estimates the exit probability and employment level, but could not identify a credible instrument that affects exits, but is uncorrelated with unobserverables in the employment equation. Thus, we estimate a conditional model that focuses on the employment decision for only surviving startups to complement the unconditional model that captures both exit and employment.

Real Estate. These sectors have growth rates that are on average 0.04 and 0.27 jobs per startup below level of Other Services.

Most of the coefficients are much larger in absolute magnitude than the coefficients from the unconditional model based on the full sample. The finding that the sector coefficients from the conditional and unconditional samples line up reasonably well suggests that differences in survivor rates are not driving the overall differences by major industry group. This is consistent with the finding from Table 3.2 that survival rates do not differ substantially across industry groups. The unconditional model includes a value of zero employees after a startup exits. Since exit rates are similar across sectors including them will simply result in a uniform shift downward in absolute magnitude in coefficients.

The regression results confirm what we report in Figure 5.1: Average job creation per startup grows very quickly among surviving startups and continues to increase over time. Controlling for macroeconomic conditions and sectors through the regressions does not change this conclusion. To put it simply, surviving startups tend to grow fast on average.

Movement between No Employment, Employment and Exit

Another important type of underlying heterogeneity in the dynamics of startups is the movement between not having employees, having employees and exits. Only a small share of all startups hire employees in the initial year, but many non-employer startups transition to hiring employees over time. Because of data limitations these movements have not been previously examined in detail.

Figure 7 displays the evolution of startups over time across the trichotomy of not hiring employees, hiring employees, and exits. The percentages of the startup cohort that hire employees, do not hire any employees and have exited in the startup year and over the seven follow-up years are reported. Eleven percent of the startup cohort hire employees in the initial year with the rate declining steadily over time. By the seventh follow-up year, 7 percent of the startup cohort hires employees. In contrast, 89 percent of the startup cohort have no employees in the initial year and by the seventh follow-up year 22 percent have no employees. The decline in the percentage of startups hiring employees or not hiring employees is primarily driven by the rapidly increasing percentage of startups that exit. By the seventh follow-up year, 71 percent of the startup cohort has exited.

At startup, there are 8.25 nonemployer firms for every employer firm. In the first follow-up year, this drops to 4.4 nonemployer firms for every employer firm and then steadily decreases for each additional year since startup. After seven follow-up years, there are 3.3 nonemployer firms for every employer firm. Naturally, these startups with employees did not all start with employees. In fact, 19 percent of startups with employees at year 7 started with no employees.

Distribution of Employment among Startups

The results displayed in Figure 6 indicate that there is rapid growth in average employment levels among surviving startups. An important question is whether the strong upward trend in average employees per surviving startup is due to a few very fast-growing survivors (“gazelles”) or to a larger number of steadily growing survivors. To investigate this question, we return to the comprehensive startup panel and examine the transition across employment size classes over time for all startups who survive to Year 1. For this analysis we look at broad employment groups in Year 1 and track how firms change employment groups by Year 5. Table 4.1 reports the Year 5 distribution (of survivors) for each Year 1 employment group, as well as the share of each Year 1 employment group that exits by Year 5. We see a significant share of firms exiting by Year 5 for each Year 1 employment group—and this is even conditional on surviving the initial shakeout between Year 0 and Year 1.[19] We also see only modest employment growth overall by firms surviving to Year 5. Firms are by far most likely to remain in the same broad employment group, and they are more likely to move to a lower employment bin than a higher one.

Despite the transition patterns in Table 4.1, we still see an upward shift in the size distribution of startups from Year 1 to Year 5. Table 4.2 shows the share of surviving startups in each of our four broad employment categories in Year 1 and Year 5. Conditioning on survival allows us to compare the distributions in the two years using the same set of firms.[20] We find that the share of startups in the two smallest employment groups decreased in Year 5, while the share in the two largest groups increased. This upward shift occurs despite what we find in Table 4.1 because although the startups that grow represent a small share of firms in their Year 1 employment group, these startups represent a relatively larger share of their Year 5 employment group. For example, while only 8.9% of startups with 1–9 employees in Year 1 transitioned to 10–99 employees in Year 5, these startups represent 36% of the firms with 10–99 employees in Year 5. And even though only 2.8% of startups with 10–99 employees in Year 1 transitioned to 100+ employees in Year 5, these startups make up 35% of the firms with 100+ employees in Year 5. These patterns are consistent with the growth dynamics evidenced in Haltiwanger et al. (2012). They show that firm growth is highly skewed with relatively few firms disproportionally contributing to overall job creation. By contrast, most firms remain small or exit. These "up or

[19] Note from Figure 5.2 that roughly 40% of startups exit by Year 1.
[20] Note from Table 5.5 that smaller startups are more likely to exit by Year 5. For this reason, including all Year 1 startups in our comparison would bias the Year 1 distribution downward.

out" dynamics appear to be reinforced here by using the universe of startups in our comprehensive startup panel. Further research into the underlying growth patterns and accompanying entrepreneurial motivation is necessary to fully understand high growth firms amongst this universe.

Together these findings tell a story of a very tumultuous life for business startups. Many firms enter every year, but many of those do not survive past the first few years. These businesses exhibit very low rates of hiring any employees. Despite the large exit rates, and given the large volume of entrants, those that do survive and hire employees contribute significantly to job creation and growth. Perhaps most dramatic is the amount of business experimentation that these numbers suggest. Understanding the motivation for these business formations, their success, and growth remain an underexplored area of research.

VI. Heterogeneity in the Definition of an Entrepreneur

In this section, we explore two broad questions about the definition of an entrepreneur. First, do non-employer startups create jobs one, two or more years later, and if so, do they make a sizeable contribution to job creation in the United States? The previous literature has not been able to answer this question because of data limitations. Second, we explore the related question of whether the low levels of job creation per entrepreneur and survival rates found above are simply due to an overly inclusive definition of entrepreneur (i.e. using the entire startup universe). If we restrict the population of startups to be less inclusive and require a stronger signal of commitment by entrepreneurs at startup, do we continue to find low levels of job creation and survival?

All of the analyses of job creation and survival dynamics thus far use the most inclusive definition of entrepreneurship possible – the universe of business startups in the U.S. economy. But there is substantial heterogeneity that exists across business types at startup. To investigate this heterogeneity, we start by focusing on the main distinction between startup types, which is whether they are non-employers or employers at the time of startup. The distinction arises primarily because of how the two types of businesses are treated in Census data. As noted above, non-employer businesses are generally tracked by PIKs and employer businesses are tracked by EINs. Using the iLBD we can identify whether the first observation in the administrative panel data is non-employer or employer. This is not possible by only using the administrative LBD panel data, which has been used in many previous studies, because it does not include the pre-employment histories of each employer firm.

Focus on Employer Startups

Previous research and statistics on entrepreneurs, startups or small businesses published by the U.S. Census Bureau, U.S. Bureau of Labor Statistics and U.S. Small Business Administration focus almost exclusively on new *employer* businesses or new *employer* establishments. As noted above, new employer businesses are identified by the first year with employees and do not look back for potential non-employer startup histories. Published numbers from these sources or numbers that can be calculated from published sources, as expected, indicate higher levels of job creation and survival rates than for the startup universe. In the well-publicized "Frequently Asked Questions about Small Business" report, the SBA reports[21] that the average number of employees per new employer firm is roughly 6. The SBA also reports that

[21] See https://www.sba.gov/sites/default/files/advocacy/SB-FAQ-2017-WEB.pdf. These statistics are derived from underlying Census or BLS data.

an average of 79 percent of new employer establishments survive one year later and roughly 50 percent of all new employer establishments survive five years or longer.

Recently, the BLS started publishing aggregate data under their "Entrepreneurship in the U.S. Economy" series based on the Business Employment Dynamics (BED) series.[22] Statistics on the number of establishments that are less than 1-year old are used to measure "entrepreneurs." Averaging over many cohort years, these data indicate that roughly 50 percent of new employer establishments survive at least five years. Job creation can also be calculated from reported BED data and indicate that new employer establishments hire an average of roughly 6 employees in the first year. Unfortunately, only the first year is available from published sources by the BLS.

The U.S. Census Bureau publishes statistics on the "Number of firm startups" based on the Business Dynamics Statistics (BDS) series. Job creation and survival rate numbers can be calculated for new employer businesses.[23] The BDS is created from the underlying Longitudinal Business Database (LBD), which includes only employer businesses. Firm startups are defined as those employer businesses with age equal to 0. Thus, in this case and with recent research analyzing the underlying data entrepreneurial activity is represented as new *employer* businesses (e.g., Decker et al. 2014). But, the focus is on new businesses and not new establishments, which is an advantage of the BDS over the BED data. Using information from multiple years and age of businesses, cohorts can be tracked over time. We calculate that new employer businesses create an average of 6.1 jobs in the first year and 4.7 jobs in the fifth year after startup.[24] Survival rates

[22] See https://www.bls.gov/bdm/entrepreneurship/entrepreneurship.htm. The BED is generated from the Quarterly Census of Employment and Wages (QCEW) database. The underlying QCEW data come from employment and total wage information covered by state and federal unemployment insurance programs.
[23] As noted in Haltiwanger, Jarmin and Miranda (2013) prior to the release of the published BDS there was no age information in publically available data leading to the perceived finding of an inverse relationship between firm size and growth in the data.
[24] For employer business data by firm age see https://www.census.gov/ces/dataproducts/bds/data_firm.html.

are 76 percent survive after one year, 65 percent survive after 2 years, and 46 percent survive after 5 years.

For comparison, a non-government source of data on startup survival rates is the Kauffman Firm Survey (KFS). The KFS follows a sample of roughly 5,000 non-employer and employer startups from the 2004 cohort. For this sample of startups, roughly 90 percent survive one year and 55 percent survive 5 years after startup (Robb and Farhat 2013). The underlying sampling frame of the KFS, however, is skewed towards including employer firms, which may also explain why these estimates of survival rates are much higher than for the universe of startups.

Contributions of Non-Employer Startups

Given the previous focus on employer startups, it is important to understand what is lost by not including non-employer startups. In particular, do non-employer startups create jobs in the years following startup, and how many jobs do they create? Figure 8 displays the number of jobs created by the average employer startup and the average non-employer startup over time. Note that non-employer startups that do eventually hire an employee are classified as employer firms in those years but they are not classified here as employer startups. Job creation per non-employer startup is substantially lower than job creation per employer startup in the follow-up years, as expected. The average employer startup creates 6.3 jobs in the first year after startup falling to 4.1 jobs seven years after startup. Seven years later, the average business that started without employees has 0.09 employees.

Although these statistics appear to indicate that employer startups generate essentially all jobs among startups, this is not the case. There are many more non-employer startups than employer startups. Non-employer startups represent roughly 90 percent of all startups. The total number of existing non-employer businesses also dwarfs the total number of existing employer businesses. Non-employer businesses represent more than three quarters of all businesses in the United States.[25]

Examining total jobs created by all employer startups vs. all non-employer startups reveals a somewhat different pattern than jobs created per entrepreneur. Figure 9 displays the total number of jobs created by employer and non-employer startups over time. The total number of jobs created by non-employer startups steadily increases relative to the total number of jobs created by employer startups over time. One year after startup, non-employer startups hire 181,000 employees. Employer startups hire 2.9 million, which is considerably higher. But, seven years after startup, we find that non-employer startups hire 319,000 employees and employer startups hire 2.0 million employees.[26] Non-employer startup total job creation increases from 6 percent of the total one year after startup to 14 percent of the total seven years after startup.

Additional Arguments for the Inclusion of Non-Employer Startups

The inclusion of non-employer startups is also important for conceptual reasons. A large part of the entrepreneurship literature takes an individual approach to analyzing entrepreneurship. For example, classic studies of the entrepreneur such as Knight (1921), Schumpeter (1934) and more recently Evans and Jovanovic (1989) do not limit the definition of entrepreneurs to include only business ventures with employees. Additionally, most previous empirical studies of entrepreneurship using a vast range of different datasets do not distinguish

[25] Data are from the BDS and Non-Employer Statistics.
[26] Comparing the jobs created by surviving non-employer and surviving employer startups over time we find that the jobs created by the two types of startups in terms of payroll per employee do not differ. Average pay per employee is very similar between non-employer and employer startups.

between entrepreneurs with or without employees (for a few examples see Evans and Leighton 1989; Holtz-Eakin, Joulfaian and Rosen 1994; Blanchflower and Oswald 1998; Fairlie 1999; Hamilton 2000; Cullen and Gordon 2007; Fairlie and Robb 2008; Lafontaine and Shaw 2016; Levine and Rubinstein 2016, 2018; Hvide and Oyer 2018, and review in Parker 2009).

Many non-employer businesses are also extremely successful. Non-employer data reported by the U.S. Census Bureau indicate that there are nearly 300,000 non-employer businesses with ½ million or more in annual revenues in 2014. The latest U.S. Census Bureau Survey of Business Owners (SBO) data indicate that there are nearly 250,000 non-employer businesses with ½ million or more in annual revenues in 2012. Part of the entrepreneurial folklore especially in high-tech is how many very successful entrepreneurs started in garages without employees.

Thus, leaving out non-employer startups may obscure a relevant dimension of early-stage entrepreneurial dynamics in the United States. In particular, it would be misleading to exclude all non-employer businesses to answer the question of how many jobs are created by the average entrepreneur. Not only does it affect the denominator in this calculation, but many non-employer startups become employer firms and make substantial contributions to job creation, sometimes several years after startup. Additionally, inclusion of non-employer startup data is important for accurately measuring survival rates and identifying the cohort year and years since startup as many young employer firms begin as non-employer startups. In fact, among firms hiring their first employee (which is the LBD definition of an employer startup), roughly 80 percent of these firms are employer startups and 20 percent are non-employer startups from previous years.

Restricting the Non-Employer Startup Population

We next examine whether the low levels of job creation per entrepreneur and survival rates found above are simply due to an overly inclusive definition of entrepreneur (i.e. using the entire startup universe). If we restrict the population of startups to be less inclusive and require a stronger signal of commitment by entrepreneurs at startup, do we continue to find low levels of job creation and survival?

There are essentially two data-driven primary classifications of business entities by the U.S. Census Bureau and four legal form subclassifications within the non-employer universe (and another subclassification within this):

1) Employer startups
2) Non-employer startups
    - 2.1) Incorporated: business granted a charter recognizing it as a separate legal entity having its own privileges and liabilities distinct from those of its members.
    - 2.2) S-Corporation: A form of corporation where the entity does not pay any federal income taxes, and its income or losses are divided among and passed to its shareholders (which are reported on their own individual income tax returns).
    - 2.3) Partnership: unincorporated business where two or more persons join to carry on a trade or business with each having a shared financial interest in the business.
    - 2.4) Sole proprietorship with EIN: unincorporated business with a sole owner.
    - 2.5) Sole proprietorship without EIN: unincorporated business with a sole owner.

Instead of excluding all non-employer startups, we want to use this information to narrow down the definition of an entrepreneur. The goal is to propose an alternative definition of entrepreneurship that might serve as a reasonable upper bound on job creation per entrepreneur and survival rates. We include all non-employer startups that are incorporated, S-corporations, partnerships, or sole proprietorships with EINs but exclude sole proprietorships without EINs. In each of these cases there is a much stronger business registration requirement than for sole proprietorships without EINs. Also, consultants and contract workers will typically show up as sole proprietorships, and never file for an EIN. These work arrangements are technically classified as business entities in the data because of treatment in the tax code, but they do not fit the theoretical concept of an entrepreneur nor the conventional view of a business owner. They receive a 1099 form from a business and do not have to do anything else. Unfortunately, however, using this approach will result in the loss of some growth-oriented sole proprietorships that initially do not file for EINs that eventually hire employees.

This approach is in line with recent work that refines the definition of the entrepreneur. For example, Guzman and Stern (2016) use characteristics at founding such as how the firm is organized (e.g., as a corporation, partnership, or LLC, and whether the company is registered in Delaware), how it is named (e.g., whether the owners name the firm eponymously after themselves), and how the idea behind the business is protected (e.g., through an early patent or trademark application) to identify high potential startups.[27] They note that although new businesses can be organized as sole proprietorships, more formal registration is a useful prerequisite for growth and includes benefits such as limited liability, tax benefits, the ability to issue and trade ownership shares, and credibility with potential customers. Levine and Rubinstein (2016, 2018) separate the incorporated and unincorporated self-employed in the CPS and NLSY to identify "entrepreneurs" and "other business owners," respectively. They demonstrate that the incorporated self-employed engage in activities requiring strong non-routine cognitive abilities, tend to be more educated, score higher on learning aptitude tests, exhibit greater self-esteem, engage in more illicit activities when young, and experience a large increase in earnings relative to wage/salary work. They also find that incorporated business owners are

[27] Their sample covers 15 states capturing roughly 50 percent of the economy.

positively selected on human capital and collateral and move with the business cycle, but unincorporated owners do not.

Figure 10 displays the total number of jobs created by startups and sole proprietorship startups over time. The total number of jobs created by this more restrictive population of startups is 3.0 million in the first year after startup. Total job creation drops to 2.3 million in the seventh follow-up year. Excluding sole proprietors without EINs results in the loss of 19,000 jobs in the first follow-up year to 56,000 jobs in the seventh follow-up year (which is 18 percent of the jobs created by all non-employer startups).

As expected, the bigger effect from excluding non-EIN sole proprietor startups is on the measure of job creation per entrepreneur. The numerator (total number of jobs) is smaller, but the real change is a much smaller denominator (total number of entrepreneurs). Figure 11 displays the number of jobs created by the average startup excluding sole proprietors without EINs and the average sole proprietor without EIN startup over time. Job creation per non-sole proprietor without EIN startup is higher than job creation per all startups in the follow-up years. The average non-sole proprietor without EIN startup creates 2.6 jobs in the first year after startup then declines to 1.9 jobs seven years later. Seven years later the average sole proprietor without EIN startup hires 0.02 employees.

We also estimate regressions for unconditional and conditional job creation that exclude sole proprietor startups. Specification 1 of Table 5 reports estimates from the unconditional model, and Specification 2 reports estimates from the model in which we condition on survival in that year. The same sets of variables and controls are included in the regressions. The sample size is now smaller and the means of the dependent variables are larger. We find many similarities between the results. The years since startup and industry coefficients generally follow the same patterns. Employment levels drop over time when using the sample of all firm-years and increase over time when using the sample of surviving firm-years. Industry patterns are generally similar after removing sole proprietor without EIN startups. See Appendix Table 3 for the full industry breakdown.

We can certainly expect estimates of job creation per startup to be higher if all non-employer startups are excluded. On the other hand, many non-employer startups grow and ultimately hire employees many years later underestimating counts of total jobs created per startup cohort. Conditioning on a subset of non-employer startups seeks to remedy these potential biases.

Survival Rates

As expected, we find that survival rates are higher among the more restrictive group of startups, but surprisingly not by much. Figure 12 displays survival rates for all startups, non-sole proprietor without EIN startups, and sole proprietor without EIN startups. Although sole proprietor without EIN startups have lower survival rates, their inclusion is not the primary reason that survival rates are low among the universe of startups. One year after startup, 76 percent of non-sole proprietor without EIN startups survive. This is higher than for the universe of startups (59 percent) and sole proprietor without EIN startups (55 percent), but the difference is not extremely large. The difference in survival rates is fairly consistent over time since startup. Five years after startup, the survival rate for non-sole proprietor without EIN startups is 45 percent, compared with a survival rate of 33 percent for the universe of startups and 31 percent for sole proprietor without EIN startups.

Specification 2 of Table 3 reports regression results excluding sole proprietor startups. The same sets of variables and controls are included in the regression. The coefficient estimates do not change substantially. For the years since startup, the coefficients in fact are very similar. We find the same pattern of a rapid decline in survival rates and tapering off after that. For industries, we continue to find only small differences in survival rates across industries. Excluding sole proprietor startups from the "entrepreneur" sample does not change the findings regarding industry differences except for a couple of exceptions.

Lower and Upper Bounds on Job Creation and Survival

The inclusion or exclusion of sole proprietor, non-employer startups is an important decision in measuring entrepreneurial job creation and survival. Instead of taking a stand on whether entrepreneurship is best measured by the universe of startups or our new non-sole proprietor definition of startups we consider the two measures as lower and upper bounds, respectively.

Table 6 presents average jobs per entrepreneur using the full universe of startups as a lower bound and the non-sole proprietor subpopulation of startups as an upper bound. At one year after startup, the lower and upper bounds for job creation per entrepreneur are 0.74 and 2.56, respectively. Five years after startup the lower to upper bound range for jobs per entrepreneur is 0.63 to 2.15.

Table 6 also presents upper and lower bounds for entrepreneurial survival rates using the full universe of startups and subpopulation of non-sole proprietor startups. One year after startup, 59 to 76 percent of entrepreneurial firms survive. This drops to 47 to 64 percent in the second year after startup and to 33 to 45 percent five years after startup.

Following a few recent studies, we restrict the definition of entrepreneur. On one hand, we lose some total jobs created, but job creation per entrepreneur is now higher. Also, survival rates are not that different, and the trend in job creation over years since startup is roughly similar.

VII. Conclusions

We create and analyze a new compilation of administrative data covering the universe of startups in the United States. Our analysis of this new U.S. startup panel dataset produces several novel findings on entrepreneurship and job creation that have important implications for policy, economic welfare, and our understanding of entrepreneurship. First, entrepreneurs make major contributions to total job creation in the United States. We find that each annual cohort of startups creates a total average of 3.0 million jobs in the first year after startup, and employs a total of 2.6 million workers five years after startup. Without the job creation contributions of startups through their first several years of existence, job creation by all businesses would be negative in most years in the United States. We also find that non-employer startups (almost entirely excluded in previous research on startups) make notable contributions to job creation. Non-employer startups create an average of 319,000 jobs seven years after startup representing one-seventh of total employment by all startups, and one-fifth of firms hiring their first employee have a non-employer history.

Second, policymakers and governments need to be careful about promoting entrepreneurship. Although startups make enormous contributions to total job creation in the U.S. economy, the contributions are mostly driven by the sheer number of new businesses created each year (approximately 4.1 million). Job creation is not high per entrepreneur – we find that using our most inclusive definition of entrepreneurship the average startup employs 0.74 workers. This suggests that we should revise the fundamental question about job creation to "How many entrepreneurs does it take to create a job?" To be sure, when we exclude sole proprietor startups to create a more restrictive definition of entrepreneurship, job creation rates per startup are higher. We find that the average startup employs 2.6 workers in the first year after startup and 2.2 workers five years later. Although the most reasonable definition of entrepreneurship is likely to lie somewhere between the two, these numbers do not imply high levels of job creation per entrepreneur, and are substantially less optimistic than 6 jobs per employer startup (SBA 2017). Therefore, public policies attempting to spur business creation need to be realistic about how many jobs can be created by new entrepreneurs. These policies generally do not target only employer startups. Furthermore, with these levels of job creation per entrepreneur, programs to spur entrepreneurship will most likely have to be low cost or primarily target high-growth entrepreneurs to end up with a positive benefit/cost ratio (at least in terms of generating jobs).

Third, in a broader sense job creation is higher because entrepreneurs create jobs for themselves and those jobs are not counted as employees in the data. Although the "entrepreneurs create jobs for themselves" argument is often made in the policy arena, the results presented here indicate that business ownership jobs do not last very long. We find extremely high exit rates among the universe of startups and even among our more restrictive population of startups. We find that only 47-64 percent of startups survive two years after startup, and 33-45 percent survive five years after startup. These survival rates are lower than the oft-cited statistic of 50 percent survival after 5 years among new employer establishments (SBA 2017). Furthermore, we find somewhat surprisingly that all major industries have high exit rates, and popular businesses such as restaurants have survival rates that are much lower. The potential loss of business owner jobs should also be included in the economic welfare calculus for entrepreneurship policies.

Fourth, average patterns conceal a substantial amount of heterogeneity in the underlying dynamics of job creation and survival by startups along several dimensions explored here. We test several hypotheses for why the average number of jobs per entrepreneur remains relatively constant over years since startup and find that although startups have extremely high exit rates, the resulting job losses from these exits are nearly offset by the opposing pattern of strong employment growth among surviving startups and the large number of non-employer startups eventually hiring employees. This rapid growth in job creation among survivors is also not driven by a few extremely successful startups, but instead by a continual upward shift in the employment size distribution. Turning to heterogeneity across startup types, we find that non-employer startups make sizeable contributions to employment several years after startup (one seventh of jobs seven years later). We argue that instead of excluding all non-employer startups, a preferred alternative "upper bound" measure of job creation and survival among entrepreneurs only excludes sole proprietor non-employer startups. Job creation rates per entrepreneur are higher, but survival rates are only slightly higher. Examining heterogeneity across industries, we find that there are a few industries that are job creators – Management, Manufacturing, and Accommodation and Food Services have average employment levels that are much higher. Further research should focus on why some startups grow rapidly and others do not, and whether government policies should target specific industries.

Finally, the analysis of the startup panel dataset sheds new light on the fundamental nature of entrepreneurship. Given the low levels of job creation per startup and high exit rates, much of entrepreneurship as measured as the universe of startups is likely to be motivated by job independence, contract/consulting work, schedule flexibility, or being part of the gig economy instead of creating innovative products, services and jobs (i.e. Schumpeterian entrepreneurship). The administrative data do not include information on individual motives for creating a business, but it does appear that entrepreneurship is more about creating *a job* than creating *jobs*, and many of the business ownership jobs created are short-term, perhaps to help smooth out underemployment or unemployment spells. The newly compiled panel dataset on the universe of U.S. startups alone, or linked with other government data sources, has much potential for future research on these questions and additional ones. Further research using these data to evaluate the cost effectiveness of the multitude of proposed and existing policies to encourage entrepreneurship is especially needed.

References

Acs, Zoltan J., Brian Headd and Hezekiah Agwara. 2009. *The Nonemployer Start-up Puzzle*, U.S. Small Business Administration, Office of Advocacy Working Paper.

Audretsch, David B. 1991. "New-firm survival and the technological regime." *Review of Economics and Statistics*: 441-450.

Audretsch, David B., and Talat Mahmood. 1995. "New firm survival: new results using a hazard function." *Review of Economics and Statistics*: 97-103.

Birch, David L. 1979. *The Job Generation Process*, Report prepared for the U.S. Department of Commerce, Economic Development Administration, Washington, D.C.

Blanchflower, David G., and Andrew J. Oswald. 1998. "What Makes an Entrepreneur?" *Journal of Labor Economics*, 16 (1), pp. 26-60.

Brown, J. David and John S. Earle. 2016. Finance and Growth at the Firm Level: Evidence from SBA Loans, *Journal of Finance*. 2016.

Brown, Charles, and James L. Medoff. 2003. "Firm age and wages." *Journal of Labor Economics* 21, no. 3: 677-697.

Carroll, Robert, Douglas Holtz-Eakin, Mark Rider, and Harvey S. Rosen. 2000. "Income taxes and entrepreneurs' use of labor." *Journal of labor Economics* 18, no. 2: 324-351.

Chatterji, Aaron, Kenneth Y. Chay and Robert W. Fairlie. 2014. "The Impact of City Contracting Set-Asides on Black Self-Employment and Employment," *Journal of Labor Economics*, 32(3): 507-561.

Chrisman, James J. 2016. Economic Impact of Small Business Development Center Counseling Activities in the United States: 2014-2015, Mississippi State University.

Cullen, Julie Berry, and Roger H. Gordon. 2007. "Taxes and entrepreneurial risk-taking: Theory and evidence for the US." *Journal of Public Economics* 91.7-8: 1479-1505.

Davis, Steve J., and John Haltiwanger. 1991. "Wage dispersion between and within US manufacturing plants, 1963-86." *Brookings Papers on Economic Activity*.

Davis, Steven J., John Haltiwanger, Ron Jarmin, C.J. Krizan, Javier Miranda, Al Nucci and Kristen Sandusky. 2007. "Measuring the Dynamics of Young and Small Businesses: Integrating the Employer and Nonemployer Universes," NBER Working Paper No. 13226.

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, no. 3: 3-24.

Dunn, Thomas A. and Douglas J. Holtz-Eakin. 2000. "Financial Capital, Human Capital, and the Transition to Self-Employment: Evidence from Intergenerational Links," *Journal of Labor Economics* 18 (2): 282-305.

Evans, David and Boyan Jovanovic. 1989. "An Estimated Model of Entrepreneurial Choice Under Liquidity Constraints." *Journal of Political Economy* 97(4):808-27.

Evans, David, and Linda Leighton. 1989. "Some Empirical Aspects of Entrepreneurship," *American Economic Review*, 79, pp. 519-535.

Fairlie, Robert W. 1999. "The Absence of the African-American Owned Business: An Analysis of the Dynamics of Self-Employment." *Journal of Labor Economics*, 17(1): 80-108.

Fairlie, Robert W., and Javier Miranda. 2017. "Taking the Leap: The Determinants of Entrepreneurs Hiring their First Employee." *Journal of Economics & Management Strategy* 26(1): 3-34.

Fairlie, Robert W., Javier Miranda, and Nikolas Zolas. 2019. "Measuring job creation, growth, and survival among the universe of start-ups in the United States using a combined start-up panel data set." *ILR Review*, 72(5): 1262-1277.

Fairlie, Robert W., and Alicia M. Robb. 2007. "Why are Black-Owned Businesses Less Successful than White-Owned Businesses: The Role of Families, Inheritances, and Business Human Capital," *Journal of Labor Economics*, 25(2): 289-323.

Fairlie, Robert W., Dean Karlan, and Jonathan Zinman. 2015. "Behind the GATE experiment: Evidence on effects of and rationales for subsidized entrepreneurship training." *American Economic Journal: Economic Policy*, 7(2): 125-61.

Garcia-Macia, Daniel, Chang-Tai Hsieh, and Peter J. Klenow. How Destructive is Innovation? No. w22953. National Bureau of Economic Research, 2016.

Glaeser, Edward L., William R. Kerr, and Giacomo AM Ponzetto. 2010. "Clusters of Entrepreneurship." *Journal of Urban Economics* 67.1: 150-168.

Glaeser, Edward L., Sari Pekkala Kerr, and William R. Kerr. 2015. "Entrepreneurship and Urban Growth: An Empirical Assessment with Historical Mines." *Review of Economics and Statistics* 97, no. 2: 498-520.

Guzman, Jorge, and Scott Stern. 2016. “The State of American Entrepreneurship: New Estimates of the Quantity and Quality of Entrepreneurship for 15 US States, 1988-2014.” MIT Working Paper.

Haltiwanger, John C., Ron S. Jarmin and Javier Miranda. 2013. “Who Creates Jobs? Small vs. Large vs. Young” *Review of Economics and Statistics*, vol. 95, issue 2, pages 347-361.

Hamilton, Barton H. "Does entrepreneurship pay? An empirical analysis of the returns to self-employment." *Journal of Political Economy* 108.3 (2000): 604-631.

Headd, Brian. 2003. "Redefining business success: Distinguishing between closure and failure." *Small Business Economics* 21(1): 51-61.

Holtz-Eakin, Douglas, David Joulfaian, and Harvey Rosen. 1994. "Entrepreneurial Decisions and Liquidity Constraints." *RAND Journal of Economics* 23: 334-347.

Huber, Laura Rosendahl, Randolph Sloof, and Mirjam Van Praag. "The effect of early entrepreneurship education: Evidence from a field experiment." *European Economic Review* 72 (2014): 76-97.

Hurst, Erik and Benjamin Wild Pugsley. 2011. "What do Small Businesses Do?," *Brookings Papers on Economic Activity*, Economic Studies Program, The Brookings Institution, vol. 43(2 (Fall)): 73-142.

Hvide, Hans K., and Paul Oyer. 2018. "Dinner Table Human Capital and Entrepreneurship." National Bureau of Economic Research, Working Paper No. w24198.

Jarmin, Ron and Javier Miranda. 2002. "The Longitudinal Business Database." CES Working Papers 02-17, Center for Economic Studies, U.S. Census Bureau.

Jovanovic, Boyan. 1982. "Selection and the Evolution of Industry," *Econometrica* 50 (3).

Katz, Lawrence F., and Alan B. Krueger. 2016. The rise and nature of alternative work arrangements in the United States, 1995-2015. No. w22667. National Bureau of Economic Research.

Kihlstrom, Richard, and Jean-Jacques Laffont. 1979. "A General Equilibrium Entrepreneurial Theory of Firm Formation Based on Risk Aversion." *Journal of Political Economy*, 87(4):719-48.

Knight, Frank. 1921. *Risk Uncertainty, and Profit.* New York: Houghton Mifflin.

Kulick, Robert, John C. Haltiwanger, Ron S. Jarmin and Javier Miranda. "High Growth Young Firms: Contribution to Job, Output and Productivity Growth" Forthcoming NBER book Measuring Entrepreneurial Businesses: Current Knowledge and Challenges, John Haltiwanger, Erik Hurst, Javier Miranda, and Antoinette Schoar, editors

Lafontaine, Francine, and Kathryn Shaw. "Serial Entrepreneurship: Learning by Doing?." *Journal of Labor Economics* 34.S2 (2016): S217-S254.

Lerner, Josh. 1999. "The Government as Venture Capitalist: The Long-Run Impact of the SBIR Program." *Journal of Business* 72.3.

Lerner, Josh. 2009. *Boulevard of Broken Dreams*. Princeton University Press.

Levine, Ross, and Yona Rubinstein. 2016. "Smart and Illicit: Who Becomes an Entrepreneur and Do they earn more?." *Quarterly Journal of Economics*.

Levine, Ross, and Yona Rubinstein. 2018. "Selection into Entrepreneurship and Self-Employment." National Bureau of Economic Research Working Paper No. 25350.

Mathematica. 2017. A Study of the Self-Employment Assistance Program: Helping Unemployed Workers Pursue Self-Employment. Final Evaluation Report for U.S. Department of Labor.

OECD. 2017. *SME and Entrepreneurship Policy Series*. Organisation for Economic Co-operation and Development Press, Paris.

OECD. 2017. *Entrepreneurship at a Glance 2017*. OECD Publishing, Paris.

Parker, Simon C. (2009). *The Economics of Entrepreneurship*. Cambridge: Cambridge University Press.

Phillips, Bruce D and Kirchhoff, Bruce A. 1989. Formation, Growth and Survival; Small Firm Dynamics in the U.S. Economy. Small Business Economics, Vol. 1, Issue 1, p. 65-74 1989.

Robb, Alicia, and Joseph Farhat. 2013. An Overview of the Kauffman Firm Survey: Results from 2011 Business Activities

Schumpeter, J.A. 1934. *The Theory of Economic Development.* Cambridge: Harvard University Press.

Tracy, Spencer L. 2011. *Accelerating Job Creation in America: The Promise of High-Impact Companies*, U.S. Small Business Administration, Office of Advocacy Working Paper.

U.S. Bureau of Labor Statistics. 2017. "Survival rates of establishments, by year started and number of years since starting, 1994–2015," Business Employment Dynamics, Entrepreneurship in the U.S. Economy, https://www.bls.gov/bdm/entrepreneurship/bdm_chart3.htm.

U.S. Census Bureau. 2017. "Definitions – Business Dynamics Statistics," https://www.census.gov/ces/dataproducts/bds/definitions.html.

U.S. Small Business Administration. 2017. "Frequently Asked Questions about Small Business," https://www.sba.gov/sites/default/files/advocacy/SB-FAQ-2017-WEB.pdf.

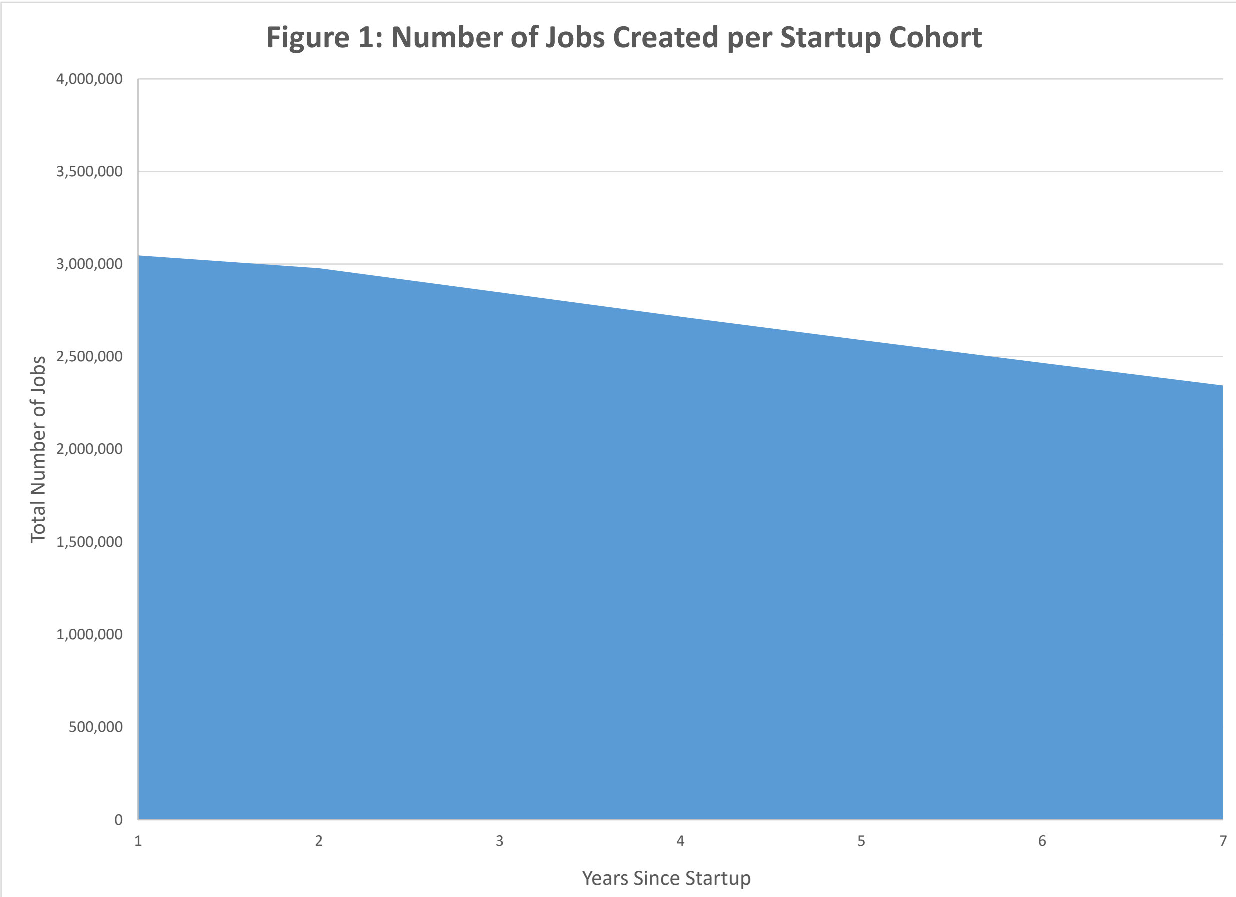
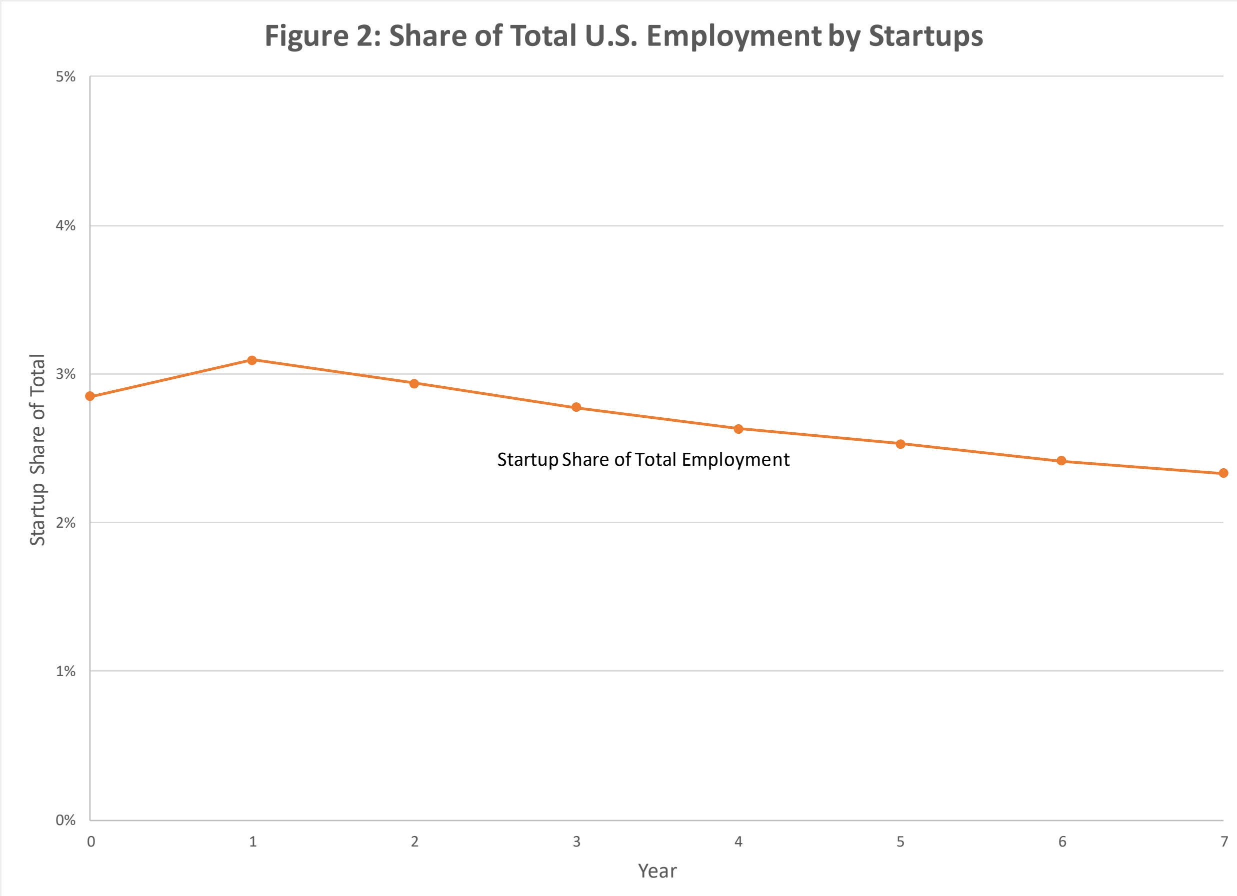


Figure 1: Number of Jobs Created per Startup Cohort
4,000,000
3,500,000
3,000,000
2,500,000
2,000,000
1,500,000
1,000,000
500,000
0
Total Number of Jobs
1
2
3
4
5
6
7
Years Since Startup



Source: Authors calculations of combined iLBD and LBD files.

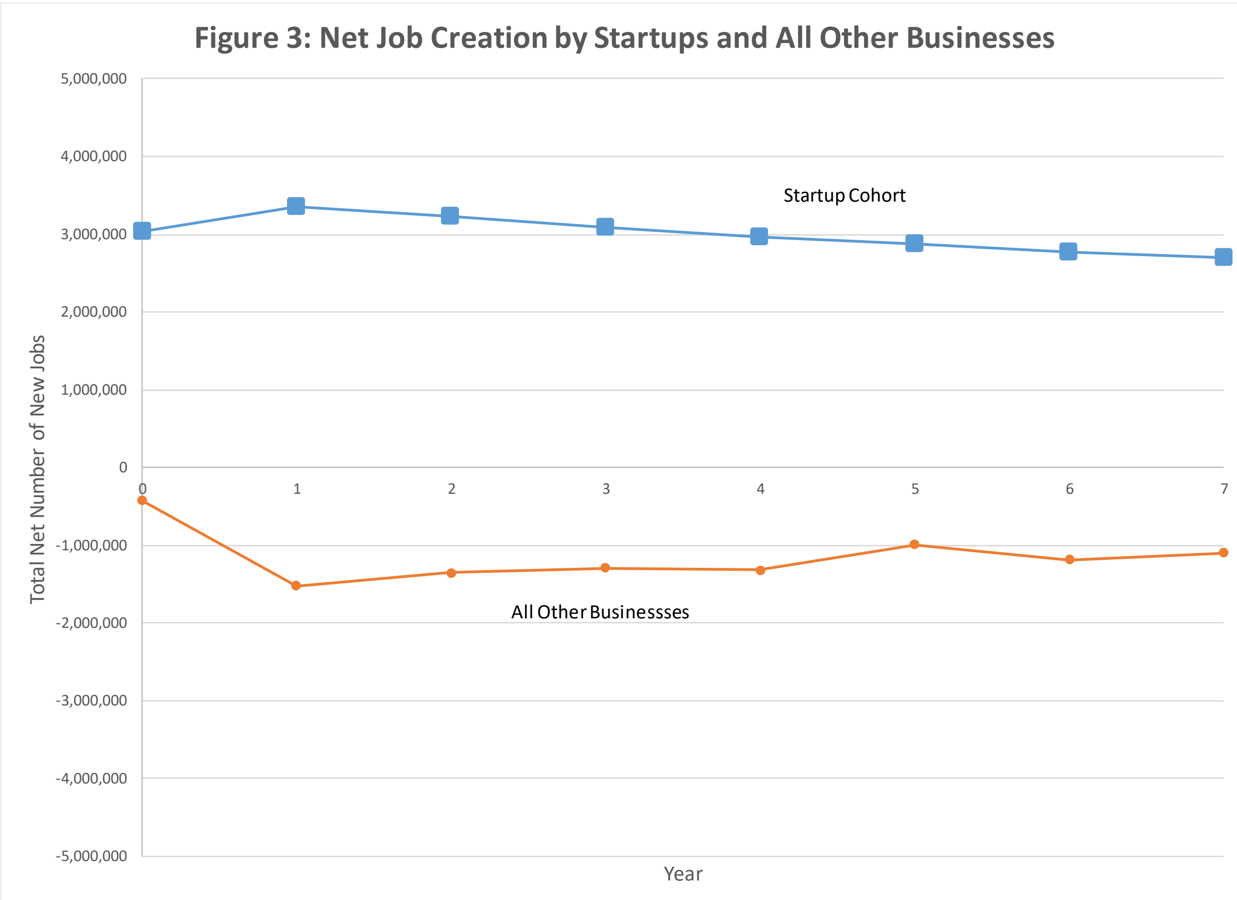


Figure 3: Net Job Creation by Startups and All Other Businesses
5,000,000
4,000,000
3,000,000
2,000,000
1,000,000
0
-1,000,000
-2,000,000
-3,000,000
-4,000,000
-5,000,000
Total Net Number of New Jobs
0
1
2
3
4
5
6
7
Year
Startup Cohort
All Other Businessses

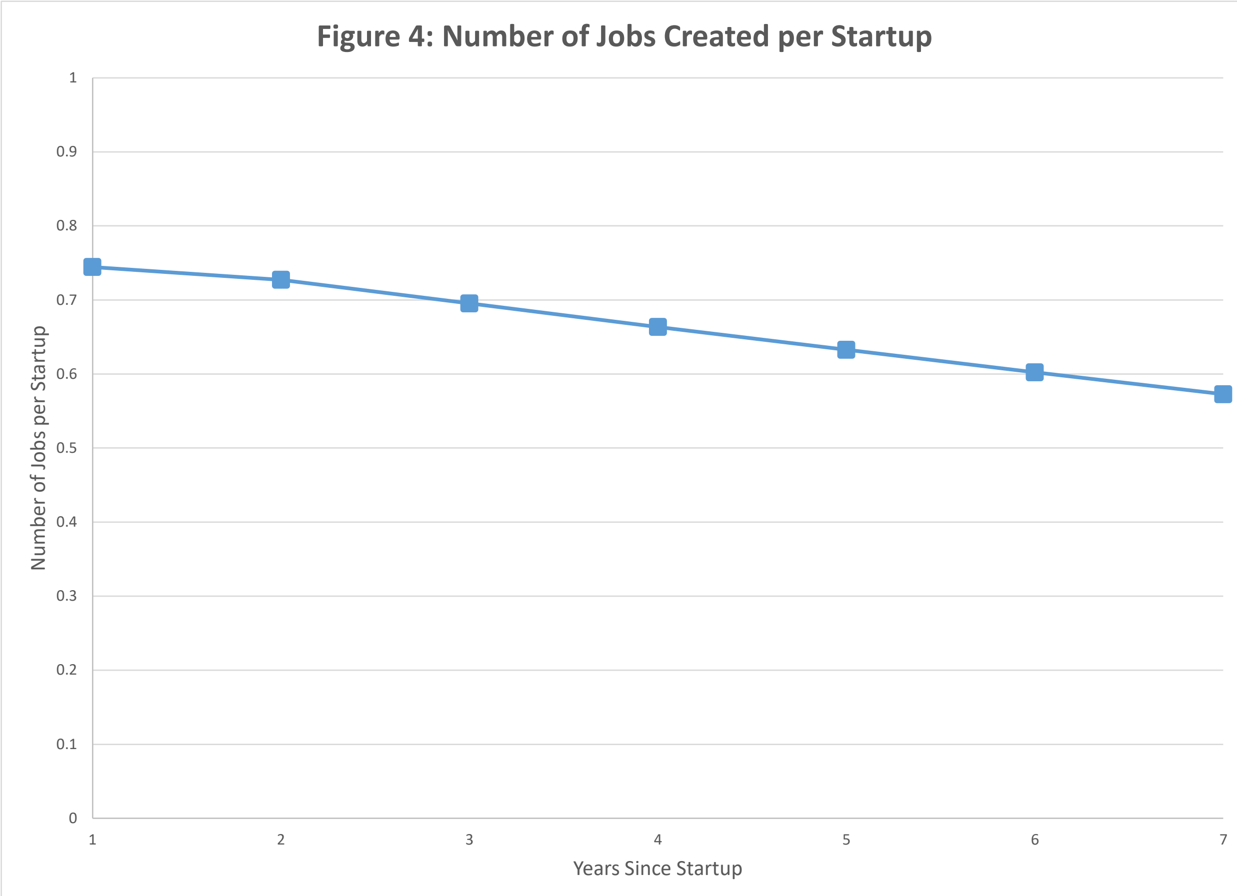


Figure 4: Number of Jobs Created per Startup
Number of Jobs per Startup
Years Since Startup
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1
2
3
4
5
6
7

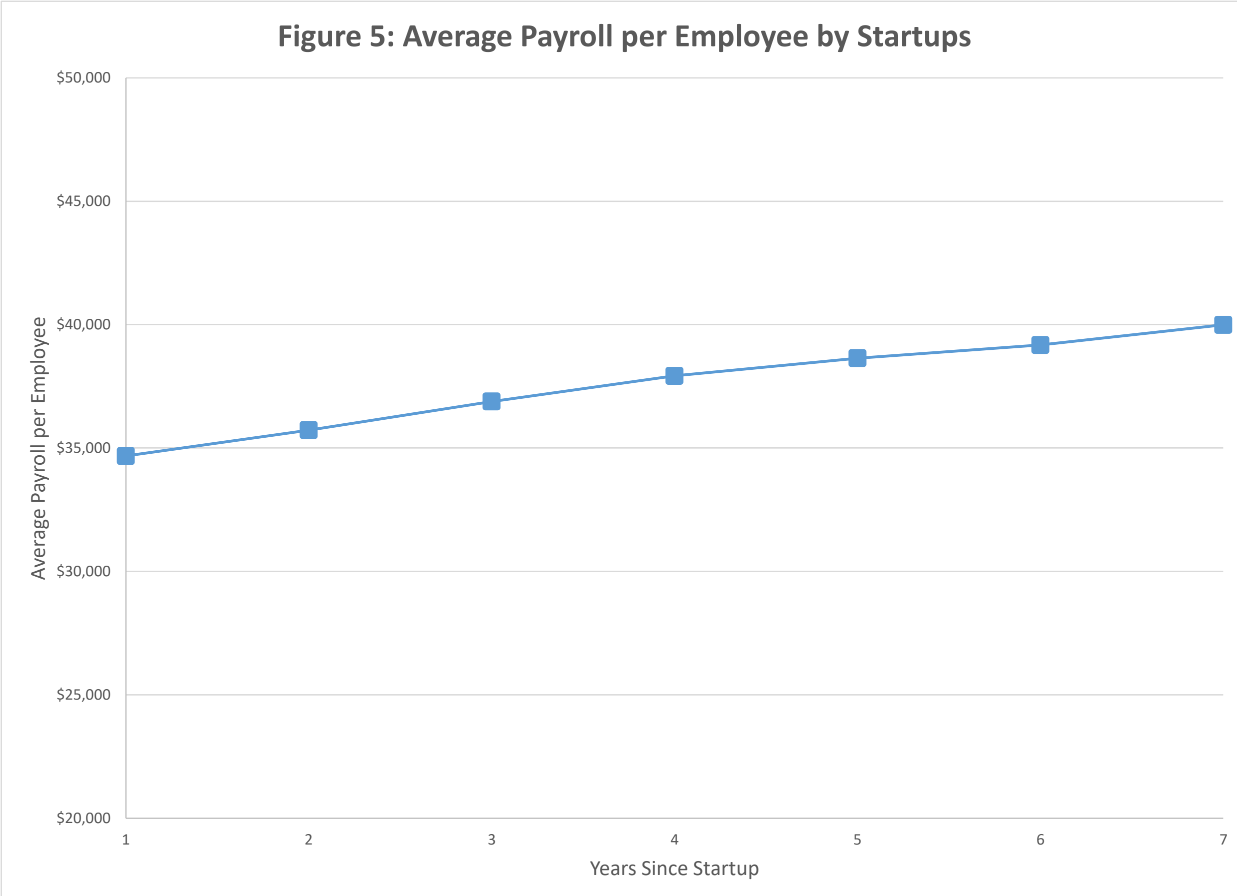


Figure 5: Average Payroll per Employee by Startups
$50,000
$45,000
$40,000
$35,000
$30,000
$25,000
$20,000
Average Payroll per Employee
1
2
3
4
5
6
7
Years Since Startup

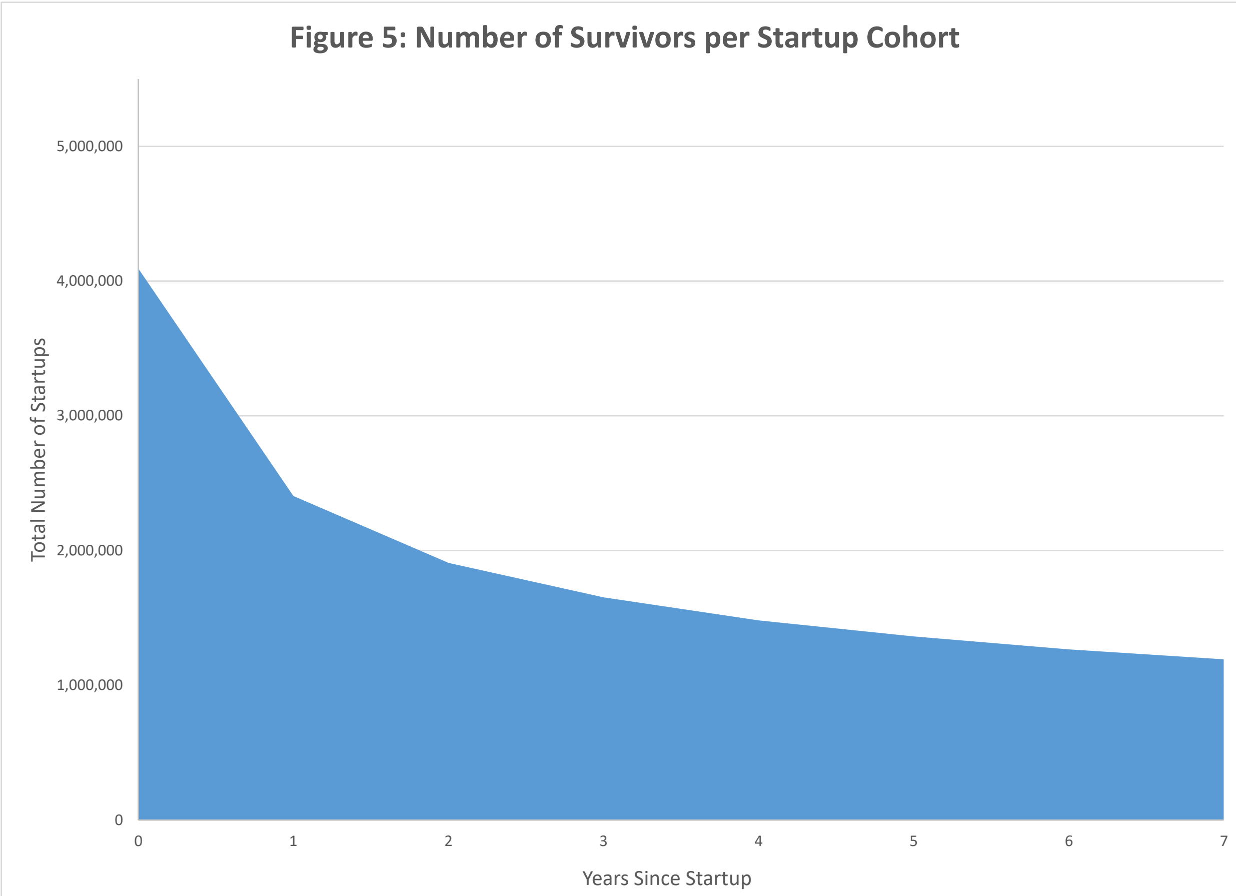


Figure 5: Number of Survivors per Startup Cohort

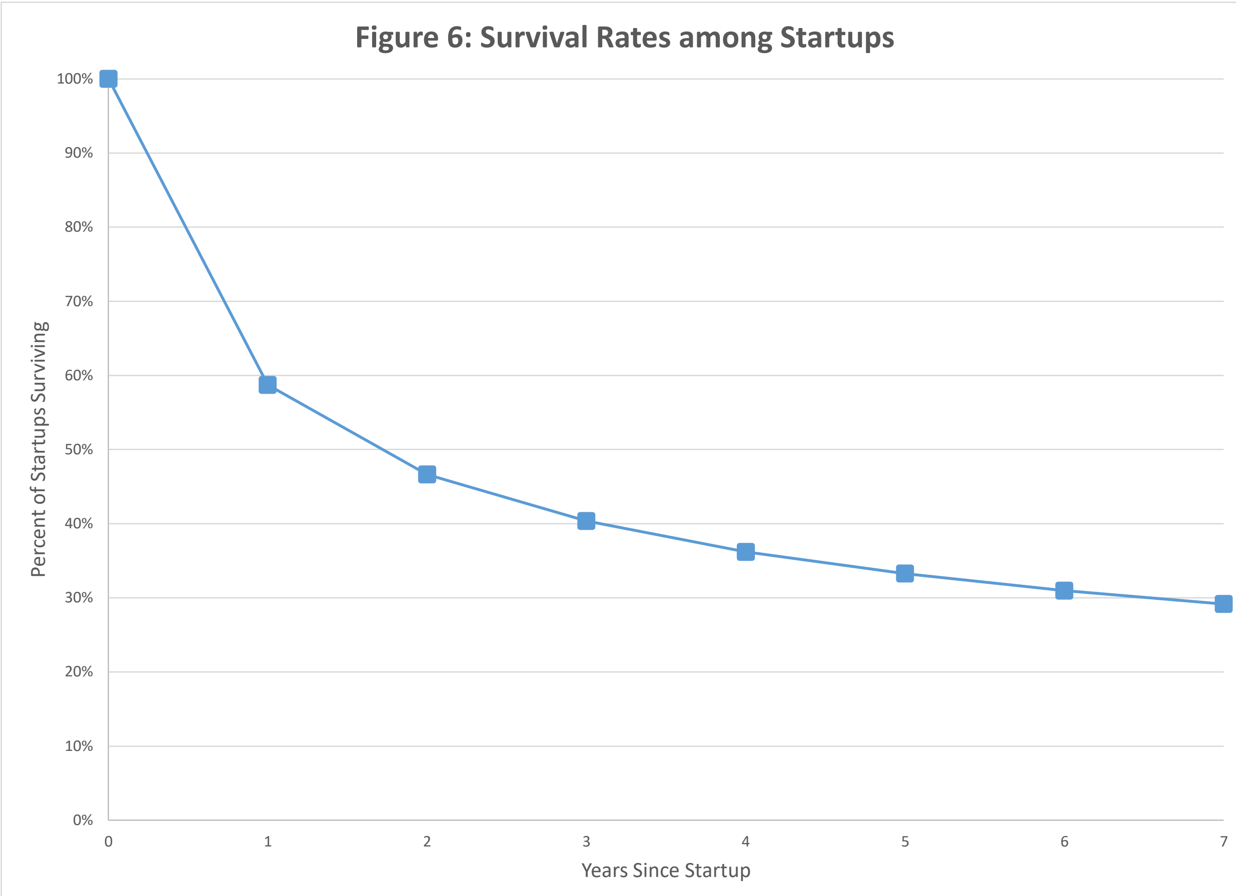


Figure 6: Survival Rates among Startups
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Percent of Startups Surviving
0
1
2
3
4
5
6
7
Years Since Startup

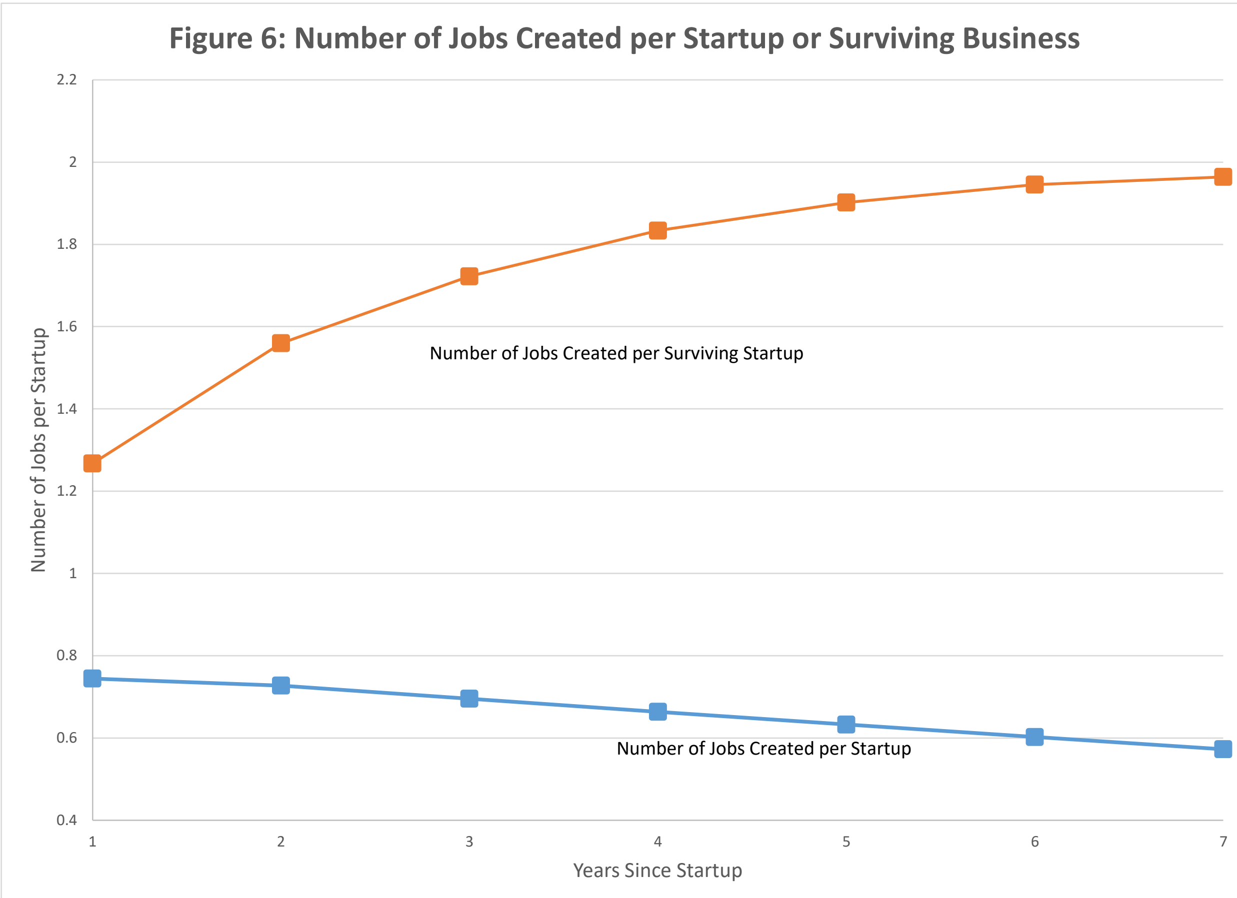


Figure 6: Number of Jobs Created per Startup or Surviving Business
Number of Jobs per Startup
2.2
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
Number of Jobs Created per Surviving Startup
Number of Jobs Created per Startup
1
2
3
4
5
6
7
Years Since Startup

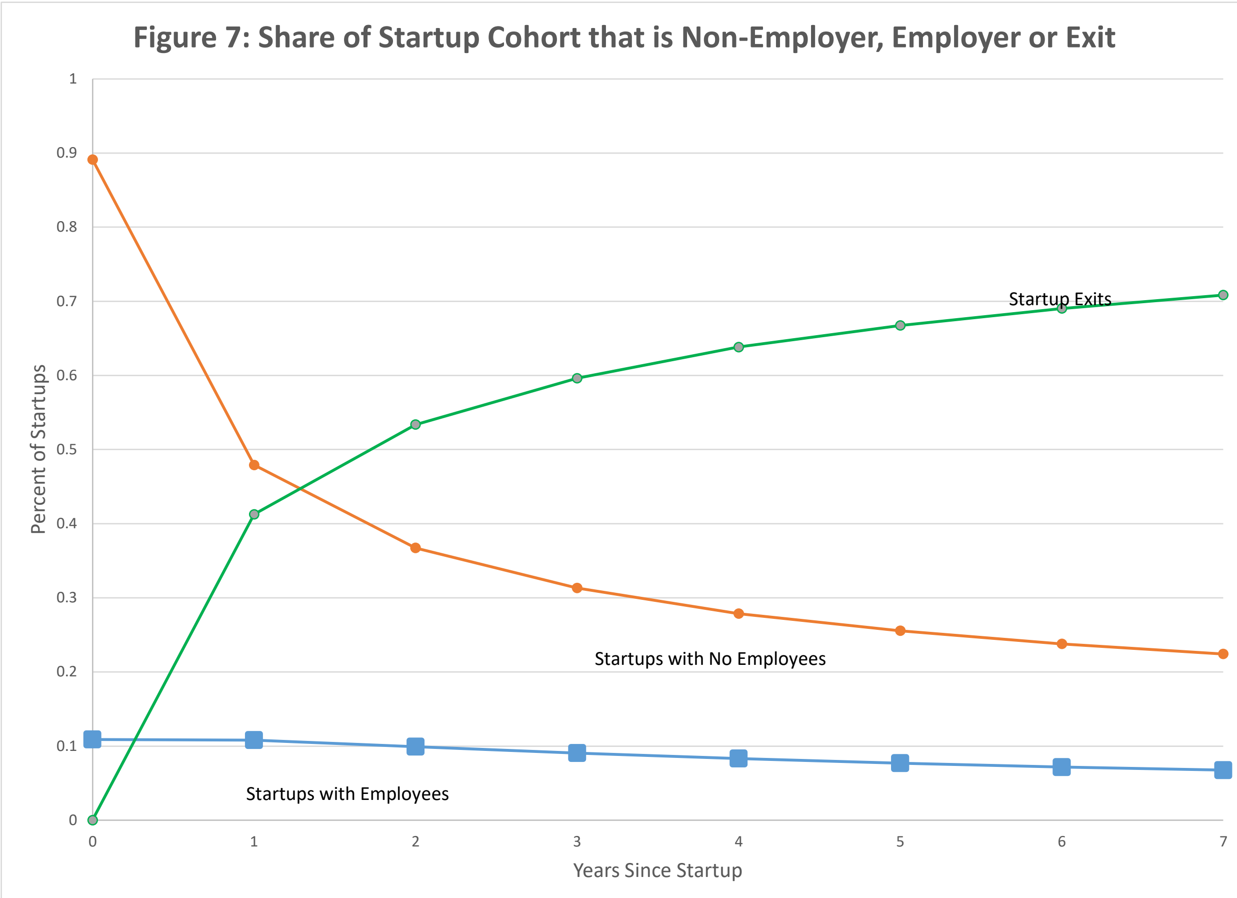
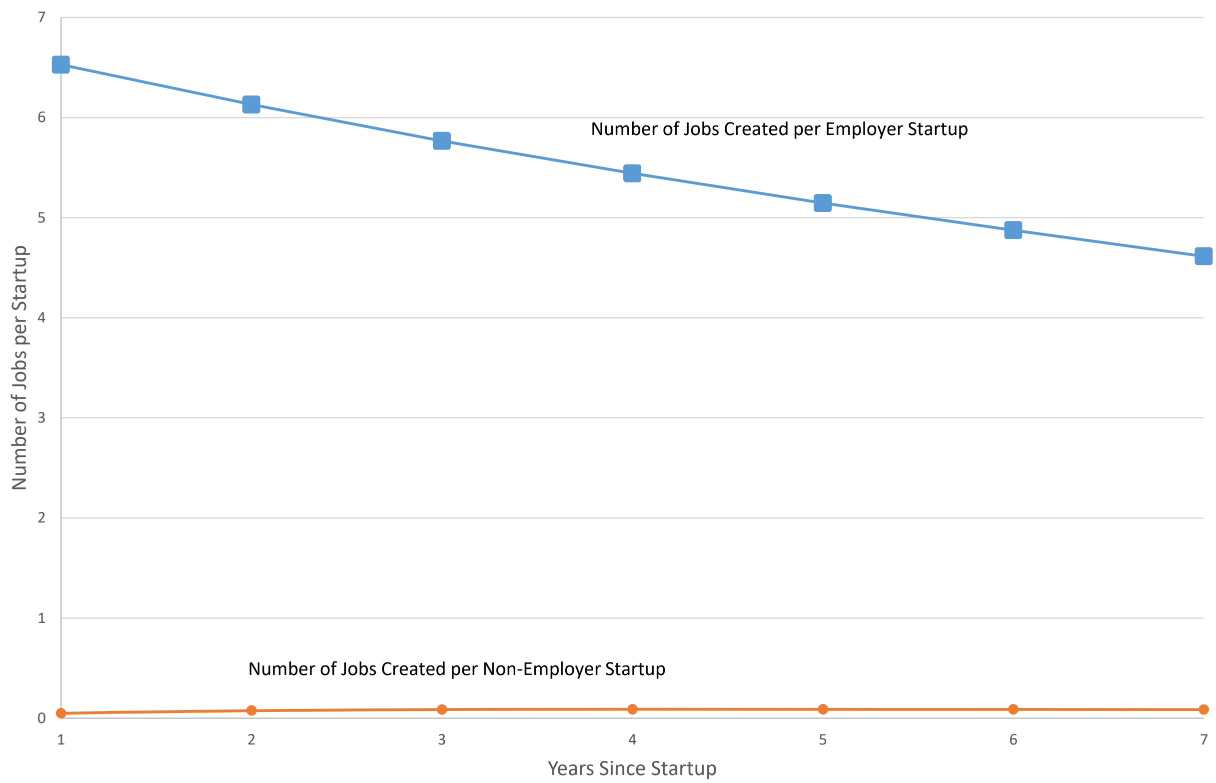


Figure 7: Share of Startup Cohort that is Non-Employer, Employer or Exit
Percent of Startups
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Startup Exits
Startups with No Employees
Startups with Employees
0
1
2
3
4
5
6
7
Years Since Startup

Figure 8: Number of Jobs Created per Startup by Employer and Non-Employer Startups



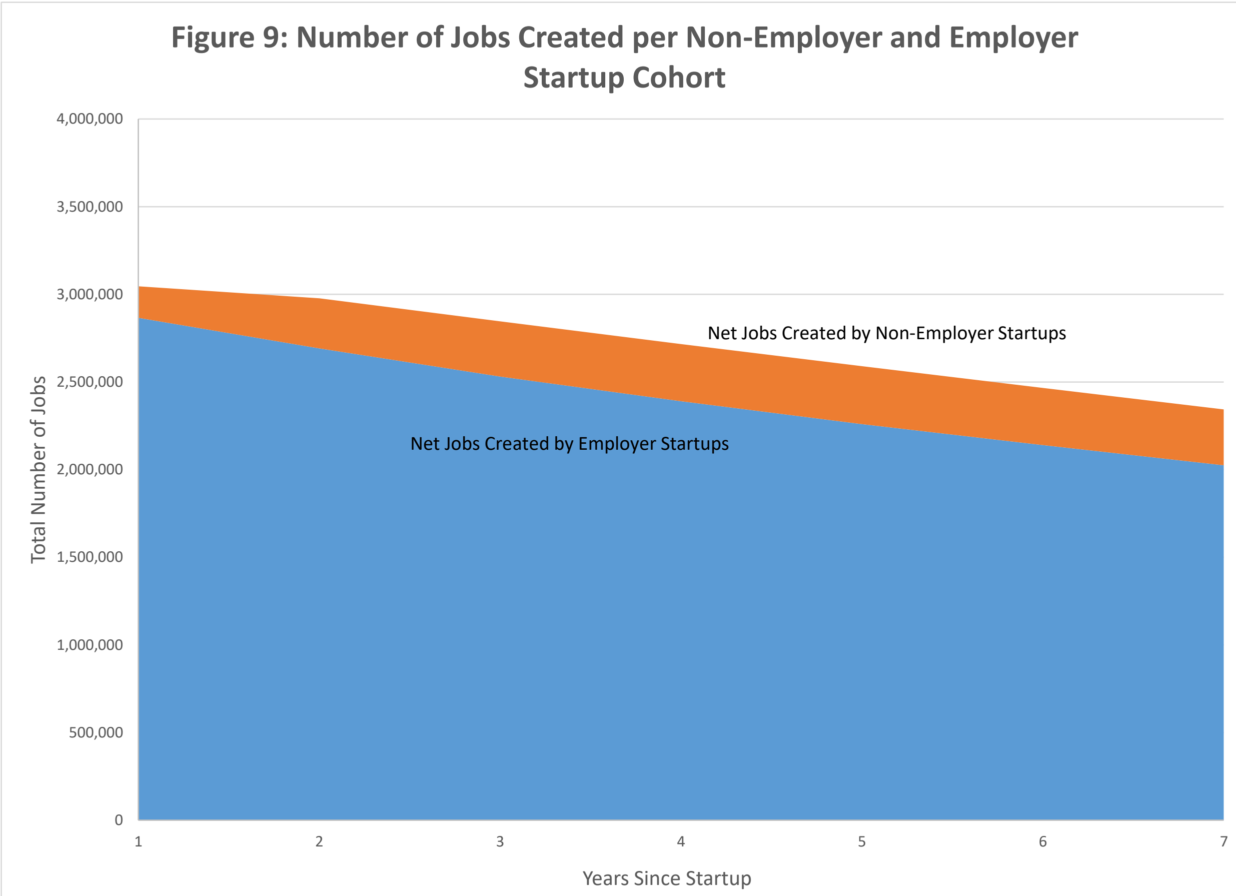


Figure 9: Number of Jobs Created per Non-Employer and Employer Startup Cohort
4,000,000
3,500,000
3,000,000
2,500,000
2,000,000
1,500,000
1,000,000
500,000
0
Total Number of Jobs
Net Jobs Created by Non-Employer Startups
Net Jobs Created by Employer Startups
1
2
3
4
5
6
7
Years Since Startup

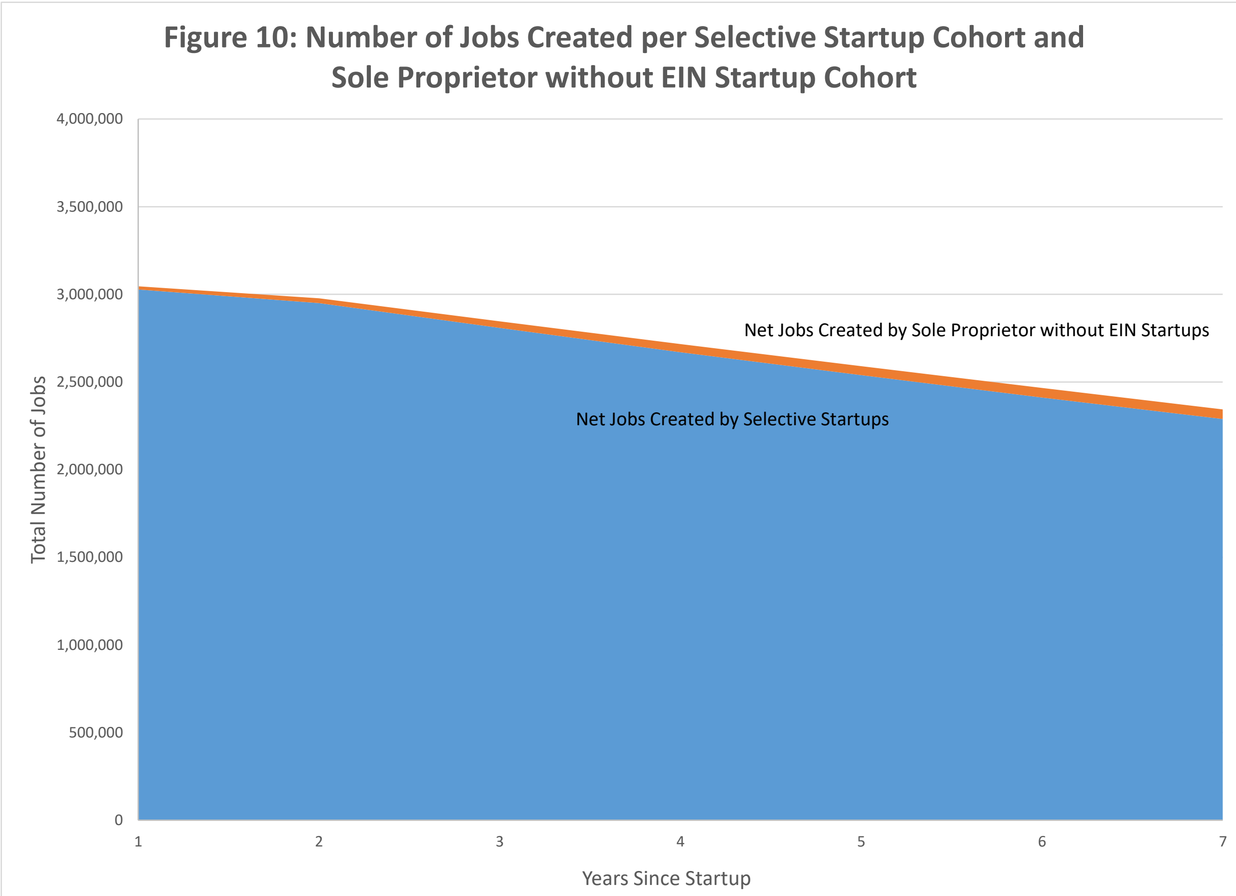


Figure 10: Number of Jobs Created per Selective Startup Cohort and Sole Proprietor without EIN Startup Cohort
4,000,000
3,500,000
3,000,000
2,500,000
2,000,000
1,500,000
1,000,000
500,000
0
Total Number of Jobs
Net Jobs Created by Sole Proprietor without EIN Startups
Net Jobs Created by Selective Startups
1
2
3
4
5
6
7
Years Since Startup

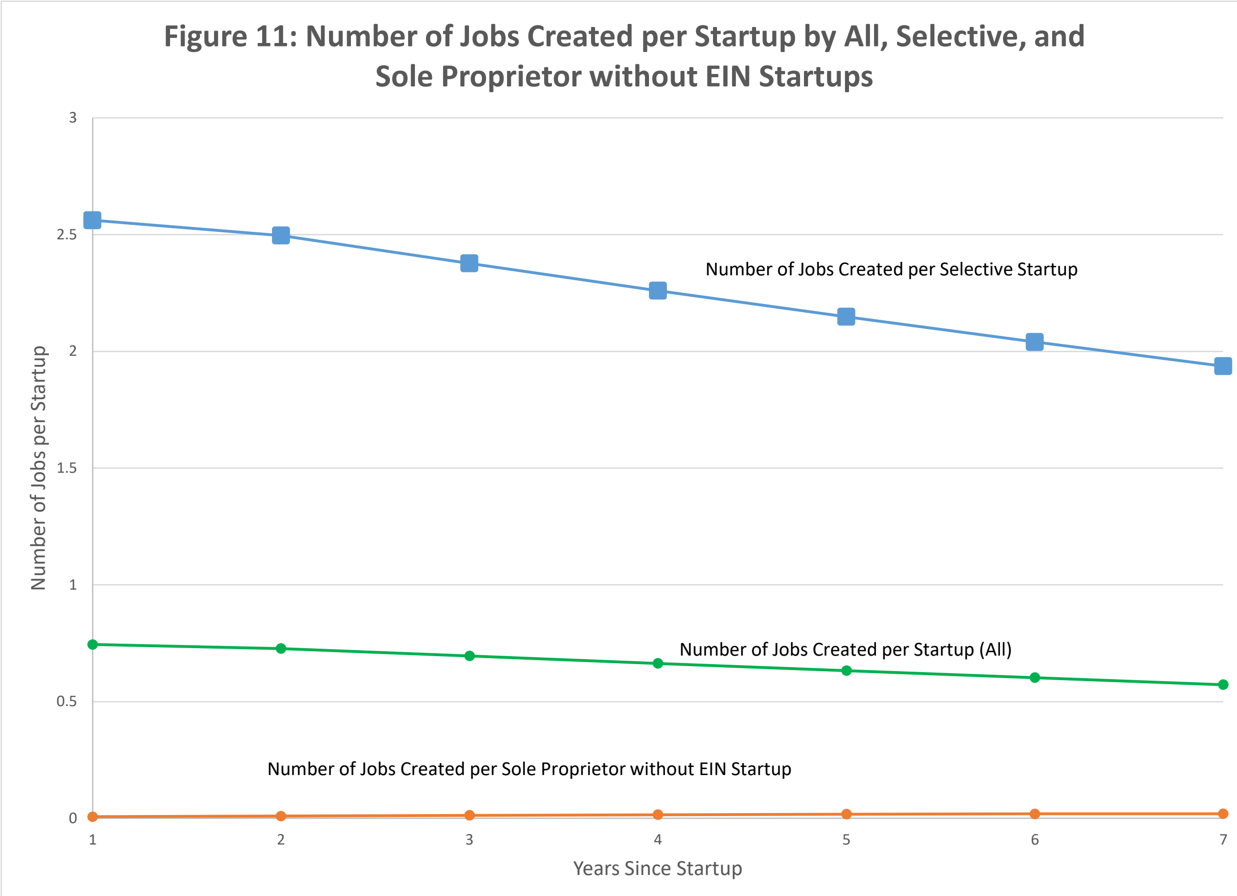


Figure 11: Number of Jobs Created per Startup by All, Selective, and Sole Proprietor without EIN Startups
Number of Jobs Created per Selective Startup
Number of Jobs Created per Startup (All)
Number of Jobs Created per Sole Proprietor without EIN Startup
Number of Jobs per Startup
Years Since Startup
3
2.5
2
1.5
1
0.5
0
1
2
3
4
5
6
7

Figure 12: Survival Rates among All, Selective, and Sole Proprietor without EIN Startups

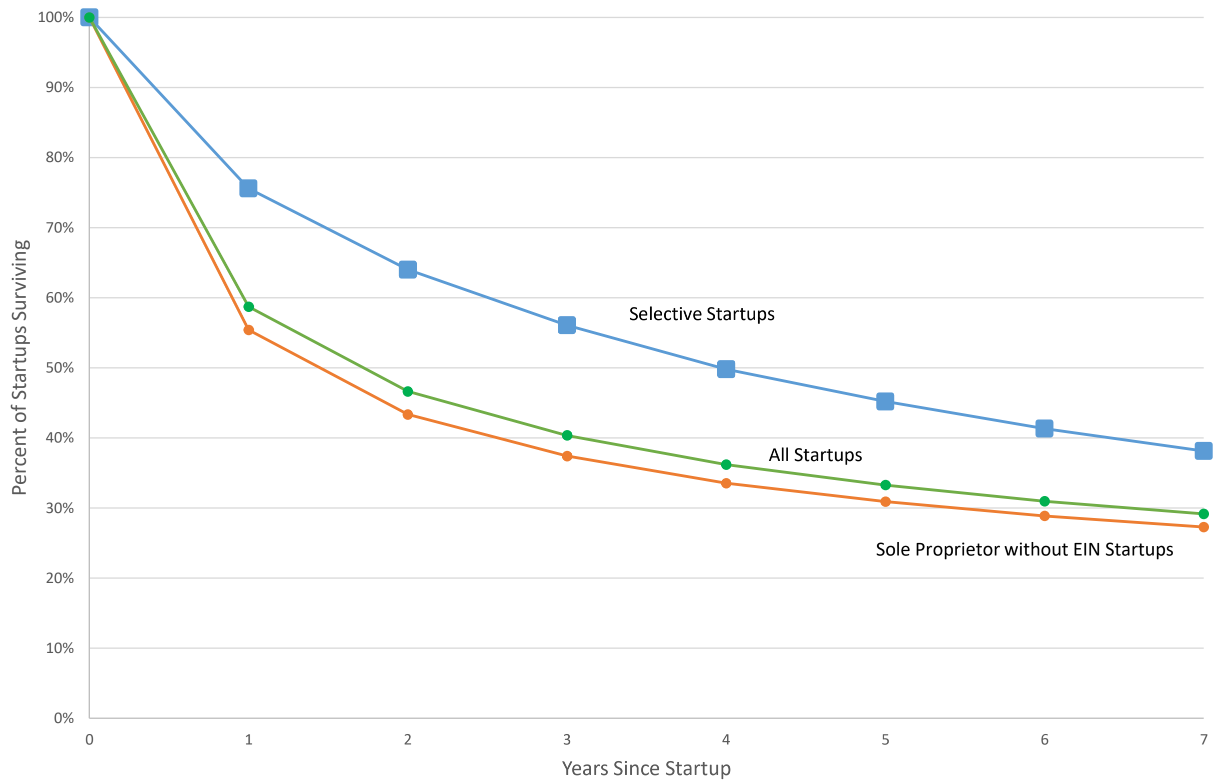


Table 1: Job Creation and Survival among Startups

Years Since Startup	Startups	Total Employment	Average Employment	Surviving Startups	Survival Rate
0	4,093,559			4,093,559	1.00
1	4,093,559	3,046,724	0.74	2,403,988	0.59
2	4,093,559	2,976,871	0.73	1,908,629	0.47
3	4,093,559	2,846,435	0.70	1,652,579	0.40
4	4,093,559	2,715,700	0.66	1,481,268	0.36
5	4,093,559	2,589,553	0.63	1,361,906	0.33
6	4,093,559	2,465,776	0.60	1,267,456	0.31
7	4,093,559	2,343,941	0.57	1,193,518	0.29

Source: Authors' calculations from the newly compiled Comprehensive Startup Panel (CPL) dataset (1995-2018). The startup year (Year 0) is the first calendar year in which the business has revenues or payroll. The following year (Year 1) captures the first complete calendar year of information on number of employees and survival in the data. Startup cohorts include 1995 to 2011, and followup years include 1996 to 2018.

Table 2: Regression for Number of Employees among Startups

	(1)	(2)
	Employment Coefficient (All Startups)	Employment Coefficient (Surviving Startups)
1 Year Since Startup	(excluded category)	(excluded category)
2 Years Since Startup	-0.019***	0.302***
	(0.001)	(0.002)
3 Years Since Startup	-0.050***	0.488***
	(0.001)	(0.003)
4 Years Since Startup	-0.082***	0.609***
	(0.001)	(0.003)
5 Years Since Startup	-0.112***	0.689***
	(0.001)	(0.004)
6 Years Since Startup	-0.143***	0.745***
	(0.001)	(0.004)
7 Years Since Startup	-0.173***	0.771***
	(0.001)	(0.004)
11-Agriculture, Forestry, Fishing and Hunting	-0.035***	-0.037***
	(0.004)	(0.008)
21-Mining	1.042***	1.829***
	(0.051)	(0.096)
22-23 Utilities & Construction	0.212***	0.516***
	(0.002)	(0.004)
31-33-Manufacturing	2.914***	5.742***
	(0.028)	(0.056)
42-Wholesale Trade	0.952***	1.789***
	(0.007)	(0.014)
44-45-Retail Trade	0.277***	0.681***
	(0.002)	(0.004)
48-49-Transportation and Warehousing	0.135***	0.246***
	(0.004)	(0.008)
51-Information	0.546***	1.352***
	(0.015)	(0.033)
52-Finance and Insurance	0.335***	0.581***
	(0.006)	(0.011)
53-Real Estate and Rental and Leasing	-0.048***	-0.270***
	(0.002)	(0.003)
54-Professional, Scientific, and Technical Services	0.144***	0.297***
	(0.002)	(0.003)
55-Management of Companies and Enterprises	2.305***	4.689***
	(0.118)	(0.230)
56-Administrative and Support and Waste Manage.	0.487***	1.237***
	(0.006)	(0.015)
61-Educational Services	0.273***	0.759***
	(0.010)	(0.024)
62-Health Care and Social Assistance	0.633***	1.485***
	(0.005)	(0.010)
71-Arts, Entertainment, and Recreation	0.063***	0.108***
	(0.004)	(0.007)
72-Accommodation and Food Services	3.198***	6.327***
	(0.011)	(0.022)
81-Other Services (except Public Administration)	(excluded category)	(excluded category)
Sample Size	556,700,000	261,200,000

Notes: Sample includes the universe of startups from the newly compiled Comprehensive Startup Panel (CSP) dataset (1995-2018). Startup cohorts include 1995 to 2011, and follow-up years include 1996 to 2018. Notes: Each observation is a firm–year combination, and all possible follow-up years are included. The dependent variable is the number of employees in that year for the startup and equals zero if the startup does not exist in that year. Standard errors are reported in parenthesis below coefficient estimates and are clustered at the firm level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table 3: Regressions for Survival Probability among Startups

	(1)	(2)
	Survival Coefficient (All Startups)	Survival Coefficient (Selective Startups)
Startup Year	(excluded category)	(excluded category)
1 Year Since Startup	-0.417 ***	-0.248 ***
	(0.0000752)	(0.000142)
2 Years Since Startup	-0.54 ***	-0.365 ***
	(0.0000752)	(0.000142)
3 Years Since Startup	-0.603 ***	-0.445 ***
	(0.0000752)	(0.000142)
4 Years Since Startup	-0.645 ***	-0.507 ***
	(0.0000752)	(0.000142)
5 Years Since Startup	-0.675 ***	-0.554 ***
	(0.0000752)	(0.000142)
6 Years Since Startup	-0.698 ***	-0.593 ***
	(0.0000752)	(0.000142)
7 Years Since Startup	-0.717 ***	-0.625 ***
	(0.0000752)	(0.000142)
11-Agriculture, Forestry, Fishing and Hunting	-0.012 ***	-0.033 ***
	(0.000199)	(0.000399)
21-Mining	0.07 ***	0.053 ***
	(0.000376)	(0.000635)
22-23 Utilities & Construction	-0.02 ***	-0.002 ***
	(0.000125)	(0.000292)
31-33-Manufacturing	0.027 ***	0.017 ***
	(0.000179)	(0.000343)
42-Wholesale Trade	0.041 ***	0.015 ***
	(0.000168)	(0.000325)
44-45-Retail Trade	-0.023 ***	-0.028 ***
	(0.000125)	(0.000292)
48-49-Transportation and Warehousing	0.017 ***	-0.013 ***
	(0.000144)	(0.000321)
51-Information	-0.041 ***	-0.072 ***
	(0.000179)	(0.000370)
52-Finance and Insurance	0.044 ***	0.012 ***
	(0.000153)	(0.000310)
53-Real Estate and Rental and Leasing	0.124 ***	0.085 ***
	(0.000133)	(0.000288)
54-Professional, Scientific, and Technical Services	0.007 ***	0.016 ***
	(0.000123)	(0.000290)
55-Management of Companies and Enterprises	-0.03 ***	-0.082 ***
	(0.000439)	(0.000541)
56-Administrative and Support and Waste Management	-0.045 ***	-0.017 ***
	(0.000128)	(0.000311)
61-Educational Services	-0.051 ***	-0.002 ***
	(0.000161)	(0.000429)
62-Health Care and Social Assistance	-0.027 ***	0.059 ***
	(0.000127)	(0.000307)
71-Arts, Entertainment, and Recreation	0.012 ***	-0.029 ***
	(0.000145)	(0.000351)
72-Accommodation and Food Services	0.027 ***	-0.002 ***
	(0.000156)	(0.000307)
81-Other Services (except Public Administration)	(excluded category)	(excluded category)
Sample Size	556,700,000	160,700,000

Notes: Sample includes the universe of startups from the newly compiled Comprehensive Startup Panel (CPL) dataset (1995-2018). Startup cohorts include 1995 to 2011, and followup years include 1996 to 2018. Notes: Each observation is a firm-year and all possible followup years are included. The dependent variable equals one if the startup survived to that year. Standard errors are reported in parenthesis below coefficient estimates and are clustered at the firm level. * p<0.05, ** p<0.01, *** p<0.001.

Table 4.1: Employment Size Transitions among Startups

Employment at Year 1	Among Surviving Startups Employment at Year 5: 0	1 to 9	10 to 99	100+	Among All Startups Status at Year 5: Survive	Exit
0	95.5%	4.1%	0.4%	0.0%	43.8%	56.2%
1 to 9	14.8%	76.3%	8.9%	0.1%	61.5%	38.5%
10 to 99	6.6%	19.8%	70.8%	2.8%	64.9%	35.1%
100+	3.8%	2.0%	18.8%	75.4%	65.6%	34.4%

Notes: Calculated from the newly compiled Comprehensive Startup Panel (CSP) dataset (1995–2018). Notes: Percentages represent shares with respect to the corresponding rows (i.e., shares of the Year 1 employment categories).

Table 4.2: Employment Size Distribution among Year 5 Surviving Startups

Employment	**Year 1**	**Year 5**
0	79.5%	78.6%
1 to 9	17.0%	16.9%
10 to 99	3.3%	4.2%
100+	0.2%	0.3%

Notes: Calculated from the newly compiled Comprehensive Startup Panel (CSP) dataset (1995–2018). Notes: Both columns represent the set of startups that survived to Year 5. The two columns show the employment distributions based on the startups' employment in each year.

Table 5: Regressions for Number of Employees using All and Surviving Startup Years (Selective Group of Startups)

	(1)	(2)
	Employment Coefficient (All Startup Years)	Employment Coefficient (Surviving Startup Years)
1 Year Since Startup	(excluded category)	(excluded category)
2 Years Since Startup	-0.067***	0.545***
	(0.003)	(0.004)
3 Years Since Startup	-0.181***	0.942***
	(0.003)	(0.006)
4 Years Since Startup	-0.297***	1.267***
	(0.004)	(0.007)
5 Years Since Startup	-0.406***	1.517***
	(0.004)	(0.009)
6 Years Since Startup	-0.514***	1.734***
	(0.004)	(0.011)
7 Years Since Startup	-0.616***	1.893***
	(0.004)	(0.011)
11-Agriculture, Forestry, Fishing and Hunting	-0.604***	-0.932***
	(0.012)	(0.022)
21-Mining	1.912***	2.781***
	(0.125)	(0.199)
22-23 Utilities & Construction	0.339***	0.603***
	(0.007)	(0.011)
31-33-Manufacturing	5.553***	9.200***
	(0.061)	(0.102)
42-Wholesale Trade	1.066***	1.726***
	(0.014)	(0.024)
44-45-Retail Trade	0.540***	1.152***
	(0.007)	(0.011)
48-49-Transportation and Warehousing	0.048***	0.153***
	(0.013)	(0.023)
51-Information	1.202***	2.781***
	(0.046)	(0.090)
52-Finance and Insurance	-0.103***	-0.236***
	(0.013)	(0.021)
53-Real Estate and Rental and Leasing	-0.939***	-1.815***
	(0.004)	(0.007)
54-Professional, Scientific, and Technical Services	0.166***	0.196***
	(0.006)	(0.011)
55-Management of Companies and Enterprises	1.714***	3.736***
	(0.140)	(0.257)
56-Administrative and Support and Waste Manage.	2.819***	5.076***
	(0.036)	(0.063)
61-Educational Services	2.761***	4.769***
	(0.080)	(0.137)
62-Health Care and Social Assistance	3.262***	4.808***
	(0.024)	(0.037)
71-Arts, Entertainment, and Recreation	0.481***	1.036***
	(0.020)	(0.036)
72-Accommodation and Food Services	4.543***	7.887***
	(0.019)	(0.032)
81-Other Services (except Public Administration)	(excluded category)	(excluded category)
Sample Size	160,700,000	94,460,000

Notes: Sample includes the selective group of startups, which excludes sole proprietor without EIN startups, from the newly compiled Comprehensive Startup Panel (CSP) dataset (1995–2018). Startup cohorts include 1995 to 2011, and follow-up years include 1996 to 2018. Notes: For Specification 1 each observation is a firm–year combination, and all possible follow-up years are included. The dependent variable is the number of employees in that year for the startup and equals zero if the startup does not exist in that year. For Specification 2 each observation is a firm–year combination, and all surviving follow-up years are included. The dependent variable is the number of employees in that year for the startup and is removed from the sample if the startup does not exist in that year. Standard errors are reported in parenthesis below coefficient estimates and are clustered at the firm level. * p<0.10, ** p<0.05, *** p<0.01.

Table 6: Range of Measures for Employment per Entrepreneur and Survival Rates

Years Since Startup	Employees per Entrepreneur		Survival Rates	
	Lower	Upper	Lower	Upper
1	0.74	2.56	59%	76%
2	0.73	2.50	47%	64%
3	0.70	2.38	40%	56%
4	0.66	2.26	36%	50%
5	0.63	2.15	33%	45%
6	0.60	2.04	31%	41%
7	0.57	1.94	29%	38%

Source: Author calculations based on combined LBD and iLBD database

Appendix Table 1: Startup Job Creation by Major Industry Group

Major Industry Group	Number of Startups	Average Employment in Year 1	Average Employment in Year 5
Agriculture, Forestry, Fishing and Hunting	51,265	0.29	0.22
Mining	10,088	1.53	1.53
Utilities & Construction	449,941	0.56	0.48
Manufacturing	70,412	3.89	3.10
Wholesale Trade	87,706	1.41	1.30
Retail Trade	426,294	0.70	0.53
Transportation and Warehousing	162,824	0.49	0.42
Information	75,353	0.88	0.79
Finance and Insurance	126,706	0.70	0.62
Real Estate and Rental and Leasing	267,471	0.27	0.22
Professional, Scientific, and Technical Services	515,706	0.46	0.44
Management of Companies and Enterprises	11,559	2.06	1.55
Administrative and Support, and Waste Mngt.	359,118	0.83	0.75
Educational Services	107,471	0.49	0.57
Health Care and Social Assistance	375,588	0.90	0.97
Arts, Entertainment, and Recreation	165,412	0.39	0.32
Accommodation and Food Services	120,059	4.82	3.25
Other Services (except Public Administration)	504,882	0.32	0.26

Appendix Table 2: Startup Survival Rates by Major Industry Group

Major Industry Group	Number of Startups	Survival Rate Year 1	Survival Rate Year 2	Survival Rate Year 3	Survival Rate Year 4	Survival Rate Year 5	Survival Rate Year 6	Survival Rate Year 7
Agriculture, Forestry, Fishing and Hunting	51,265	0.59	0.48	0.42	0.38	0.34	0.32	0.29
Mining	10,088	0.67	0.60	0.58	0.52	0.48	0.45	0.42
Utilities & Construction	449,941	0.57	0.45	0.39	0.35	0.32	0.30	0.28
Manufacturing	70,412	0.66	0.54	0.47	0.42	0.38	0.35	0.32
Wholesale Trade	87,706	0.70	0.57	0.50	0.44	0.40	0.36	0.33
Retail Trade	426,294	0.62	0.48	0.40	0.35	0.31	0.28	0.26
Transportation and Warehousing	162,824	0.63	0.51	0.45	0.42	0.39	0.38	0.37
Information	75,353	0.57	0.42	0.35	0.30	0.28	0.25	0.23
Finance and Insurance	126,706	0.68	0.56	0.50	0.44	0.40	0.37	0.35
Real Estate and Rental and Leasing	267,471	0.73	0.64	0.59	0.54	0.51	0.48	0.46
Professional, Scientific, and Technical Services	515,706	0.62	0.49	0.44	0.39	0.36	0.34	0.32
Management of Companies and Enterprises	11,559	0.61	0.39	0.22	0.19	0.17	0.14	0.13
Administrative and Support, and Waste Mngt.	359,118	0.55	0.42	0.35	0.32	0.29	0.27	0.26
Educational Services	107,471	0.53	0.40	0.33	0.30	0.28	0.27	0.26
Health Care and Social Assistance	375,588	0.58	0.45	0.38	0.34	0.31	0.29	0.27
Arts, Entertainment, and Recreation	165,412	0.60	0.49	0.44	0.40	0.37	0.35	0.33
Accommodation and Food Services	120,059	0.68	0.54	0.45	0.40	0.35	0.32	0.29
Other Services (except Public Administration)	504,882	0.59	0.48	0.43	0.39	0.36	0.34	0.31
Not elsewhere specified	205,706	0.20	0.07	0.04	0.03	0.02	0.02	0.02

Appendix Table 3: Startup Average Number of Employees by Major Industry Group (Selective Definition)

Major Industry Group	Number of Selective Startups	Average Employees in Year 1	Average Employees in Year 5
Agriculture, Forestry, Fishing and Hunting	16,676	0.88	0.67
Mining	4,088	3.77	3.69
Utilities & Construction	127,000	1.97	1.66
Manufacturing	32,706	8.32	6.60
Wholesale Trade	44,294	2.78	2.53
Retail Trade	126,294	2.36	1.76
Transportation and Warehousing	48,118	1.64	1.40
Information	24,588	2.66	2.39
Finance and Insurance	60,824	1.44	1.26
Real Estate and Rental and Leasing	161,000	0.44	0.36
Professional, Scientific, and Technical Services	142,176	1.65	1.55
Management of Companies and Enterprises	9,441	2.52	1.87
Administrative and Support, and Waste Mngt.	63,235	4.66	4.17
Educational Services	13,353	3.90	4.45
Health Care and Social Assistance	72,647	4.65	4.91
Arts, Entertainment, and Recreation	31,294	2.06	1.66
Accommodation and Food Services	69,882	8.25	5.51
Other Services (except Public Administration)	99,000	1.60	1.28
Not elsewhere specified	35,118	0.25	0.26
Total Number of Selective Startups	1,181,735		

Sample excludes sole proprietor without EIN startups.

Appendix Table 4: Startup Survival Rates by Major Industry Group (Selective Definition)

Major Industry Group	Number of Selective Startups	Survival Rate Year 1	Survival Rate Year 2	Survival Rate Year 3	Survival Rate Year 4	Survival Rate Year 5	Survival Rate Year 6	Survival Rate Year 7
Agriculture, Forestry, Fishing and Hunting	16,676	0.70	0.60	0.54	0.48	0.43	0.39	0.36
Mining	4,088	0.82	0.75	0.71	0.63	0.58	0.53	0.48
Utilities & Construction	127,000	0.76	0.64	0.56	0.49	0.44	0.40	0.37
Manufacturing	32,706	0.79	0.68	0.59	0.52	0.47	0.42	0.39
Wholesale Trade	44,294	0.81	0.70	0.61	0.53	0.48	0.43	0.39
Retail Trade	126,294	0.76	0.62	0.52	0.45	0.40	0.35	0.32
Transportation and Warehousing	48,118	0.77	0.65	0.56	0.50	0.45	0.41	0.38
Information	24,588	0.68	0.53	0.43	0.37	0.33	0.29	0.26
Finance and Insurance	60,824	0.79	0.70	0.63	0.54	0.49	0.44	0.40
Real Estate and Rental and Leasing	161,000	0.81	0.73	0.67	0.62	0.58	0.55	0.52
Professional, Scientific, and Technical Services	142,176	0.78	0.67	0.59	0.52	0.48	0.44	0.40
Management of Companies and Enterprises	9,441	0.65	0.43	0.27	0.23	0.20	0.17	0.15
Administrative and Support, and Waste Mngt.	63,235	0.74	0.62	0.53	0.47	0.43	0.39	0.36
Educational Services	13,353	0.73	0.62	0.54	0.49	0.45	0.42	0.40
Health Care and Social Assistance	72,647	0.79	0.69	0.62	0.57	0.53	0.49	0.46
Arts, Entertainment, and Recreation	31,294	0.71	0.57	0.48	0.42	0.38	0.35	0.32
Accommodation and Food Services	69,882	0.82	0.67	0.57	0.49	0.43	0.38	0.34
Other Services (except Public Administration)	99,000	0.75	0.63	0.55	0.50	0.45	0.41	0.38
Not elsewhere specified	35,118	0.23	0.11	0.06	0.04	0.03	0.03	0.02

Sample excludes sole proprietor without EIN startups.