Who works for startups? The relation between firm age, employee age, and growth

Paige Ouimet, Rebecca Zarutskie

* University of North Carolina, Chapel Hill, USA
b Federal Reserve Board, Mailstop 97, 20th and C Streets, NW, Washington, DC 20551, USA

1. Introduction

Young firms are often associated with an up or out dynamic. Young firms have high failure rates, but, conditional on surviving, young firms exhibit higher average growth rates as compared to older firms. These differential dynamics between young and old firms raise important questions. What is it about young firms that makes them unique?
Do firm characteristics which differentiate between young and older firms also distinguish young firms that will grow rapidly from those that will not?

In this paper, we focus on one specific firm characteristic: the employee workforce. Labor and human capital are important components to production, especially in high-tech and innovative industries where startup activity abounds. This paper documents a number of new facts showing how employees differ at young firms, yielding insight into the characteristics and joint dynamics of young firms and young workers.

Using over a decade's worth of firm-level data from the US Census Bureau, we find that young firms employ relatively more young workers. Around 27% of employees in firms aged one to five years are between 25 and 34 years old, and over 70% are under the age of 45. In contrast, in established firms that have been in existence for 20 years or more, fewer than 18% of employees are between the ages of 25 and 34, while almost half are over the age of 45. We find similar results when we control in a regression framework for firm size, industry, geography, and time. Furthermore, we document results consistent with causality. Following a plausibly exogenous increase in young workers, we observe an increase in new firm creation.

Young employees may be more likely to match with young firms for several reasons. Given younger employees will, on average, have had more recent education, they may possess more current technical skills. Building a workforce with such characteristics can be especially critical to young firms, especially those developing new products or establishing new methods of production. In addition, young employees may be a better fit for young firms due to the fact that younger employees are likely to be relatively more risk tolerant. Greater risk tolerance can make young employees more willing to bear the labor income and human capital risk of working for a young firm or to take on riskier projects within the firm.

In addition, young firms will employ more workers who have recently completed a job search by nature of being new. Young workers are likely to switch jobs early on in their careers as they acquire job- and task-specific skills and learn about their own skills and productivity (e.g., Johnson, 1978; Topel and Ward, 1992). Thus, to the extent young workers make up a higher percentage of workers who recently completed job searches, they will be more likely to work at young firms, strictly as a function of the joint dynamics of young firms and young workers. Finally, young firms may disproportionately employ young workers due to assortative matching of workers and firms based on productivity or quality of workers and firms. To the extent that younger firms are, on average, less productive than older firms and younger workers are, on average, less productive than older workers, assortative matching would imply that less productive (young) workers would match to less productive (young) firms.

Each of these four mechanisms may explain part of the positive relation between firm age and employee age in the data. The intent of our analysis is not to identify one unique driver of the relation or even to provide a complete accounting of all the potential underlying mechanisms. Instead, we show how the evidence is consistent with three non-mutually exclusive mechanisms related to the unique skills, greater risk tolerance, and joint dynamics of young firms and young workers.

We find that the young firm, young employee relation is moderated but still significant when we limit the sample to new hires, a set of workers who have all recently undergone a recent job search. The reduction in the magnitude of the effect is consistent with the joint dynamics mechanism. Moreover, the fact that the relation still holds with new hires suggests multiple mechanisms are responsible for the employee age-firm age relation. Considering wages of young firms and young employees, we find that young firms pay lower wages, on average, a fact consistent with the existing literature. However, interesting patterns emerge when we investigate by employee age. Young employees in young firms earn higher wages than young employees in older firms. Assuming wages proxy for human capital, this result is consistent with the unique skills mechanism and does not directly support the assortative matching mechanism.

Using firm outcomes, we document several more interesting correlations involving employee age. First, firms that are created with a larger share of young employees are more likely to subsequently raise venture capital (VC) financing. Given findings in Hellmann and Puri (2000) that VC investors select more innovative firms, this result is consistent with young employees (those with recently acquired skills especially valuable for innovation) matching to more innovative startups. Second, young firms which employ relatively more young workers subsequently experience higher growth rates (conditional on survival) and, with some qualifications, higher failure rates. These results are consistent with the unique skills and greater risk tolerance of young workers mechanisms.

Finally, we ask whether the rate of new firm creation is affected by the supply of young workers. If young employees are important for young firm growth due to their unique skills or attributes, we should expect that when more young employees are available, entrepreneurs find it easier to start and grow young companies, especially in more innovative industries. Using historical demographic information on the relative ratio of youth in a state as a predictor for the ratio of younger to older workers ten or 20 years later, we argue that a causal relationship exists between the supply of young workers and the rate of new firm creation, especially in high-tech industries, where innovation is greater. These results suggest that the supply of young workers, in addition to the supply of financial capital, is an important factor in the creation and growth of new firms.

Our study contributes to the literature on what drives new firm creation and growth. A large focus of this literature has been on understanding the role of financial market development and structure. We explore the role...
of labor markets and how the relative supply of young workers can impact firm creation and growth. Our study is related to recent work by Lazear (2005), who examines the human capital traits of entrepreneurs and Doms, Lewis, and Robb (2010), who show that there is more new firm creation in regions where the local labor force is more educated.2

Our results also contribute to the labor and organizational economics literatures in documenting the strong positive association between employee age and firm age and the narrower wage spread between older and younger workers at young firms. Previous studies have explored the relation between firm size and wages (e.g., Brown and Medoff, 1989) and between firm age and wages (e.g., Brown and Medoff, 2003), but none have examined the role of employee age in explaining the relation between firm size and age with wages. Moreover, our results are applicable to the organization economics literature which suggests that firm hierarchies might be flatter in young, entrepreneurial firms (e.g., Rajan and Zingales, 2001).

The remainder of the paper proceeds as follows. Section 2 reviews the relevant literature and considers four mechanisms that could underlie the positive relation between firm age and employee age and their associated empirical predictions. Section 3 describes the data. Section 4 examines the relation between firm age and employee age. Section 5 considers evidence on the mechanisms underlying the relation. Section 6 discusses sample selection issues. Section 7 provides evidence of a causal relation between the supply of young workers and new firm creation. Section 8 concludes.

2. Possible mechanisms underlying the relation between firm age and employee age

In this section, we briefly review the related literature on firm age, employee age, and firm and career dynamics. We then consider four mechanisms suggested by this body of work that may underlie the positive relation between firm age and employee age observed in the data.

2.1. Firm age

Young firms exhibit different characteristics compared to older firms. Young firms are characterized by an up or out dynamic. For example, Dune, Roberts, and Samuelson (1989) show US manufacturing plant failure rates decline with age as do the growth rates of non-failing plants. Using more recent data covering all industries in the US, Haltiwanger, Jarmin, and Miranda (2012) also find that young firms have a high exit rate and, conditional on surviving, young firms exhibit higher growth rates as compared to older firms, controlling for size. These empirical patterns are consistent with models in which new firms enter with new business ideas or products and learn over time whether they can capture market share from incumbent firms (e.g., Jovanovic, 1982; Bahk and Gort, 1993).

Young firms are more likely to face constraints or high costs when accessing external capital, an empirical fact documented in many studies (e.g., Petersen and Rajan, 1994; Hadlock and Pierce, 2010). Such constraints can make young firms more sensitive and vulnerable to oil price and credit market shocks, as shown in Davis and Haltiwanger (2001),3 and economic downturns, as shown in Fort, Haltiwanger, Jarmin, and Miranda (2013). The greater sensitivity of young firms to financial market conditions may also lead to some of the differences in characteristics and dynamics of young and old firms as shown in the theoretical setting of Cooley and Quadrini (2001).

Young firms are often credited with a disproportionate share of innovation, relative to their older peers. Using productivity growth as a proxy for innovation, Huergo and Jaumandreu (2004) find that innovation is highest at young firms and diminishes over time. Darby and Zucker (2003) describe industry dynamics where breakthrough innovation is concentrated in a few young firms. Christiansen (1997) shows through numerous case studies how younger firms are more likely to capitalize on disruptive innovations because of differences in their organizational structures compared to older firms.

2.2. Employee age

Likewise, young workers display different behavior and characteristics compared to older workers. A large literature documenting the wage profiles of workers shows that wages tend to increase over the majority of a worker’s lifetime (e.g., Ben-Porath, 1967) with a particularly sharp increase in the first ten years of an employee’s career (e.g., Topel and Ward, 1992). This upward-sloping wage profile can reflect a variety of differences between younger and older workers and in how firms choose to compensate their employees. Younger workers may accumulate firm-specific experience and skills over time on the job (e.g., Becker, 1962). Johnson (1978) argues young workers have uncertainty as to their abilities and tastes—uncertainty that is not resolved until after sufficient job experience. As such, young workers are likely to try several jobs until they find a good match following which, presumably, wages will increase. Indeed, Topel and Ward (1992) find that in the first ten years of a career, the average worker changes jobs seven times (accounting for approximately two-thirds of total lifetime job changes) and Bjelland et al. (2011) document higher employer-to-employer turnover among younger workers. Alternatively, this rapid increase in wages in the early stage of a worker’s career has also been justified as evidence of a compensation plan where higher future wages discourage greater productivity and shirking today (e.g., Lazear, 1981; Akerlof and Katz, 1989).

2.3. Mechanisms

In this section, we posit four non-mutually exclusive mechanisms which may explain, at least in part, the

---

2 Our study is also related to the literature about how the characteristics and human capital of firm founders and top management teams relate to the characteristics and performance of the firms they join (e.g., Hamilton, 2000; Hurst and Lusardi, 2004; Gompers, Lerner, and Scharfstein, 2005; Beckman, Burton, and O’Reilly, 2007; Graham, Harvey, and Puri, 2013; and Kaplan, Klebanov, and Sorensen, 2012).

3 Age in Davis and Haltiwanger (2001) is measured at the plant level.
positive relation between firm age and employee age we observe in the data. For each mechanism, we also describe testable empirical predictions. Given the interconnected nature of these mechanisms, it is not surprising that several of the mechanisms yield similar predictions.

2.3.1. Skills

Since young employees are more likely to have recently completed their education, they may possess more current technical skills which allow them to have more innovative ideas or be able to better adapt to new environments. Building a workforce with such characteristics can be especially critical to young firms, which may be developing new products or establishing new methods of production (e.g., Christiansen, 1997).

If unique skills are a key driver of the employee age-firm age relation, then we should find that young workers at young firms have higher human capital, on average, relative to young workers at older firms. In the absence of a direct measure of human capital, we proxy with employee wages and explore whether young employees in young firms receive greater compensation relative to young employees in older firms. Furthermore, if recently acquired skills are relatively more valuable for innovation, then we should find that the tendency of young firms to hire younger employees is stronger at more innovative firms. Finally, all else equal, the skills mechanism predicts young firms with more young employees will be more successful, leading to lower firm failure and higher firm growth rates.

2.3.2. Risk tolerance

Job loss can be extremely costly for workers and young firms are more likely to fail. Moreover, young firms are more sensitive to macroeconomic shocks, and firm failures during economic downturns are likely to be especially costly to workers as alternative employment options will be most difficult to find during these periods. Gibbons and Katz (1991) find that workers affected by plant closures experienced extended unemployment and lower average wages following the event. Sullivan and von Wachter (2009) document increased risk of mortality following plant closures and estimate a high-tenure male worker (2009) document increased risk of mortality following plant closures and estimate a high-tenure male worker displaced at age 40 loses between 1 and 1.5 years of life.

As such, young firms may attract relatively more young workers due to the fact that young workers may be more risk tolerant. A number of studies in the psychology literature have found that younger people tend to be more risk tolerant (e.g., Vroom and Pahl, 1971; Hensley, 1977). The economics literature has also argued that younger people may be less risk averse when it comes to portfolio choice (e.g., Bakshi and Chen, 1994; Bodie, Merton, and Samuelson, 1992). Greater risk tolerance may make young employees more willing to bear the labor income and human capital risk of working for a young firm. Moreover, young workers may suffer less following displacement, consistent with Von Wachter and Bender (2006).

Furthermore, greater risk tolerance could mean that young workers will be more likely to select riskier projects or tasks within the firm once they are hired, leading to both higher firm growth and higher probability of firm failures. If more risk-tolerant younger employees select into riskier firms or choose riskier projects once employed, then we should observe more volatile outcomes at firms with more young employees.

2.3.3. Joint dynamics of young firms and young workers

Young firms employ workers who have recently completed a job search by nature of being new. For example, a three-year old firm can only employ workers who changed jobs in the last three years. Young workers are more likely to switch jobs early on in their careers as they acquire job-specific and task-specific skills and learn about their own skills and productivity relative to other workers and across jobs (e.g., Johnson, 1978; Topel and Ward, 1992). Thus, to the extent young workers make up a higher percentage of workers who recently completed job searches, they will be more likely to be employed at young firms. Moreover, Jovanovic (1982) and Bahk and Gort (1993) argue that firms learn about their relative strengths and weaknesses over time. With this information, young firms then adapt, changing themselves to maximize future profits. This suggests that young firms experience greater rates of employee turnover, as compared to more established firms, a prediction supported in Haltiwanger, Jarmin, and Miranda (2012). The presence of relatively greater churn at young firms will predict shorter employment tenures at these firms, strengthening the above prediction of more young workers at young firms.

Given this prediction depends critically on the assumption that young workers are relatively more likely to have engaged in a recent job search, a joint dynamics mechanism predicts a significant dampening of the young firm, young employee relation when looking at a sample of employees who have all recently completed a job search. However, after controlling for different frequencies in job searches, a learning-by-doing dynamic will still predict more young workers at young firms. Young workers joining the workforce learn about their own relative strengths and weaknesses over time, as in Johnson (1978). Older employees are more likely to know their skill sets. Given young firms are also learning and changing over time, older employees may be less inclined to work at these dynamic young firms. On the other hand, young employees who do not know their specific strengths and are looking to try different jobs may be more likely to join young firms, especially young firms that anticipate relatively greater learning and subsequent change. Jovanovic (1982) describes the outcomes from learning as either high growth or exit. As such, we expect young firms which attract more young workers to be associated with greater growth rates (for surviving firms) and greater exit rates.
2.3.4. Assortative matching based on firm and worker productivity or quality

Oi and Idson (1999a, 1999b) argue that young firms are less productive, on average, as compared to more established firms. If young employees are also less productive, then assortative matching on firm and worker productivity would suggest that more young employees would match to young firms. If assortative matching is a key driver of the positive relation between firm age and employee age, then we should expect to see the young people who join young firms are of relatively lower average quality, as compared to the young employees who join older firms. If wages are a proxy for worker quality, then this should translate into relatively lower average wages for young employees in young firms compared to young employees in older firms. Likewise, we should expect to see the young firms with the greatest number of young employees are of relatively lower average quality, as proxied by lower growth and higher failure rates.

3. Data

We use four primary data sources in the analysis. We use data from the US Census Bureau's Longitudinal Employer-Household Dynamics (LEHD) program to obtain information on the ages and wages of employees. We use data from the US Census Bureau's Longitudinal Business Database (LBD) to obtain information on the industry, age, and geography of the firms for which the employees in the LEHD work. We use Compustat to obtain additional information on the publicly traded firms in the sample. Finally, we use data from SDC Thomson VentureXpert and DowJones VentureSource to obtain information on which firms in the LBD receive venture capital financing.

3.1. Longitudinal Employer-Household Dynamics data

LEHD data are collected from the unemployment insurance records of states participating in the program. Data start in 1992 for several states and coverage of states increases over time. By 2004, 27 states in the US are included in the LEHD data.5 The LEHD data track employees who work for firms in the participating states on whom unemployment insurance taxes are paid.6

We use the Quarterly Workforce Indicators (QWI), which aggregate worker-level information in the LEHD to the business establishment level.7 The QWI data provide information on the count and total payroll for employees hired and separated each quarter. This information is reported for all employees and by age groups. Age groups are reported in ten-year intervals, e.g., age 25–34, age 35–44, etc. Total payroll includes regular salaries and all bonuses and commissions, as well as stock options and other equity compensation in some states.8 Firms in the QWI are identified by their state employer identification numbers (SEINs). Information on the physical address, industry, and federal tax employer identification number (EIN) of each business establishment is also recorded in the QWI. We annualize the QWI data by summing measures of flows, such as new hires and wages, over each quarter of a given year, and adjust wages to be in constant year 2005 dollars.

3.2. Longitudinal Business Database

The LBD is a panel data set that tracks all employer US business establishments beginning in 1975.9 The database is formed by linking years of the standard statistical establishment list (SSEL), a register of business establishments, maintained by the Internal Revenue Service of the US Treasury Department. The LBD links the business establishments contained in the SSEL over time and assigns each a unique identifier as well as a firm-level identifier that allows researchers to aggregate information to the firm level. The LBD contains information on the physical location, industry, total employment, and payroll for each business establishment.10

We use the LBD to track the business establishments of firms that are included in the QWI over our sample period 1992–2004. We measure firm age as time from first entry in the LBD, i.e., the year in which the firm hires its first employee. If a firm has multiple establishments, we take the age of the oldest establishment as the age of the firm, similar to the approach taken in Haltiwanger, Jarmin, and Miranda (2012). We can also observe the years in which an establishment exits the LBD. This allows us to identify firm shut downs. We classify a firm as exiting when all of its establishments leave the LBD, i.e., the firms’ employment goes to zero.

We link business establishments in the QWI to the business establishments in the LBD using the Business Register Bridge. These files match business establishments across the two databases using federal EIN, industry, state, and county of the establishments. We then aggregate to the firm level using the firm-level identifier in the LBD. In Appendix A, we describe our matching methodology and the characteristics of the resulting sample in more detail.

3.3. Compustat, SDC Thomson, and Dow Jones data

We link information from Compustat to the Census data using the internal Census Compustat/SSEL crosswalk.

---


6 See Abowd et al. (2006) for a more detailed description of the program and the underlying data sets that it generates.

7 A business establishment is part of a firm defined by having a particular geographic location. For example, a law firm with an office in San Francisco and an office in Los Angeles would have two business establishments. Likewise, a manufacturing firm with three different plants operating in different locations, e.g., two in Illinois and one in Wisconsin, would have three business establishments.

8 See http://www.bls.gov/cew/cewfaq.htm for additional details.

9 Data in the LBD come from a snapshot of all establishments in the US taken once a year. As such, firms which exit within the first year of life may not be included in our sample. All of our results and conclusions apply only to those firms which we can observe. Haltiwanger, Jarmin, and Miranda (2012) argue such short-lived entrants account for a small fraction of total startups.

10 For a more detailed description of the LBD, see Jarmin and Miranda (2002).
This crosswalk assigns firms in the LBD to the firm-level data in Compustat using information on EIN and location of the business establishments. Compustat contains information from publicly traded firms’ financial statements.

To link information from SDC Thomson (VentureXpert) and DowJones (VentureSource) on venture capital financings, we employ the crosswalk developed by Puri and Zarutskie (2012), which employs a name and address matching algorithm to link to firms in the LBD. Specifically, we identify firms in the LBD as VC-financed if they can be matched to firms contained in VentureXpert or VentureSource and received VC financing over our sample period.

4. The relation between firm age and employee age

We first present the main fact in our paper—that younger firms employ more young workers—before turning to an examination of the mechanisms that may drive this relation. We examine the relation between firm age and employee age for our entire sample, which is dominated by smaller private firms, as well as for a subsample of public firms.

4.1. All firms

Table 1, Panel A reports the nonparametric relation between employee age and firm age in the full sample of 4,374,025 firms. The rows in Table 1 correspond to age categories for employees; the columns correspond to age categories for firms. The final column, column 6, reports the average percentages for firms of all ages.

The QWI groups employees into age categories covering ten years, beginning at age 14, and then groups employees aged 65 and older into one category. We consider the following employee age categories: between 25 and 34, between 35 and 44, between 45 and 54, and 55 or older. We collapse the upper distribution of ages to aged 55 or older for brevity, but find similar results when we consider the categories 55–64 and 65–99 separately. We do not consider employees younger than 25 throughout our analysis since many of these employees are still completing their educations and, thus, working part-time or in temporary positions. To the extent that firms (employees) have different preferences regarding part-time and temporary workers (employment), this sample will be biased. Each cell reports the average percentage of employees in a given age category for firms in a given age category.

The striking fact that emerges from Table 1, Panel A is that younger firms disproportionately employ younger workers. Of employees in firms aged one to five years old, 26.9% are 25–34 years old, as reported in column 1. Furthermore, employees between 25 and 45 years old represent over half of the workforce at these young firms, totaling 54.9% of employees. The percentage of employees in the younger age categories falls steadily as we move across the columns and firm age increases. For firms older than 20 years in column 5, employees aged 25–34 years account for only 17.5% of employment, and employees younger than 45 for only 42.2% of employment.12

4.2. Public firms

Table 1, Panel B reports the relation between firm age and employee age for the sample of 9,120 publicly held firms in the data. We examine public firms separately to see whether the employee age-firm age relation in the full population of firms is mainly driven by smaller firms with limited access to external financial markets. For the sample of publicly held firms only, we define firm age as time from initial public offering (IPO). To become publicly held, a firm must grow to a certain size as well as meet other regulatory criteria for its equity to be publicly traded. Therefore, the sample of publicly traded firms is tilted towards relatively more established and larger firms vis-à-vis the sample of all firms.13

For firms that have gone public one to five years ago, 35.4% of employees are aged 25–34. For public firms that have gone public over 20 years ago, only 23.8% of employees are aged 25–34. In contrast, for public firms that have gone public one to five years ago, 16.4% of employees are aged 45–54. For firms which have been public for over 20 years, this percentage rises to 23.8%. There is no strong pattern between employees aged 35–44 and firm age. Thus, the employee-age firm-age relation is generally robust to the very different sample of larger public firms. In the next section, we test for statistical significance in a regression framework.

4.3. Controlling for firm size, industry, geography, and time

Table 2 tests whether the firm age-employee age patterns shown in Table 1 are robust to controls for firm size as well as year, industry, and geography fixed effects. Table 2 reports estimates from ordinary least squares (OLS) regressions of the fraction of firm employees in an age category on dummy variables for the four youngest firm age categories. The dummy variable capturing firms over 20 years is excluded from the regression. The coefficients on the firm age dummy variables should be interpreted as the difference between that age category and firms over 20 years. The regression also includes controls for firm size, measured by the lagged natural logarithm of one plus total firm employees and industry (four-digit SIC code), and state and year fixed effects for the full sample of public firms over the sample period.

11 Results for employees under age 25 are available in the Internet Appendix.

12 If we also include employees younger than 25 in the calculations, 70% of employees are younger than 45 in firms aged one to five years. For firms over 20 years, only 50% are younger than 45.

13 Summary statistics for the two samples are available in Tables A3 and A4.

14 For the full sample, we find decreasing fractions of employees in the two youngest age groups by firm age and increasing fractions of employees in the two oldest age groups by firm age. For public firms, we continue to observe the same pattern with the youngest and two oldest employee age groups; however, we find a relatively flat relation between employees age 35–44 and firm age. The lack of robustness in the 35–44 age group may indicate the inflection point between younger and older workers resides in this age group.
private and public firms.\textsuperscript{15,16} Panel A of Table 2 reports the results for all firms; Panel B reports the results for the sample of public firms.

Table 2, Panel A shows that after controlling for firm size, industry, geography, and time, the relation between employee age and firm age is strengthened. The fraction of employees age 25–34 is 9 percentage points higher in firms aged one to five compared to firms older than 20 years. This is an economically meaningful difference given that, on average, 23% of a firm's workforce is between 25 and 34 years of age, as reported in Table 1. The percentage of employees aged 25–34 steadily declines as we move up the firm age categories. We also note that larger firms employ slightly more young workers.

Table 2, Panel B shows that young public firms employ nearly 10% more employees aged 25–34, compared to the sample of firms over 20 years. There is a strong increase in the percentage of employees aged 45–54 and aged 55 or older as firms age. However, consistent with the raw data in Table 1, there is no strong firm age pattern in the data for employees aged 35–44. The relation between the percentage of employees aged 35–44 is two percentage points lower at firms aged one to five vis-a-vis firms over 20 years.

Overall, we show that young firms disproportionately employ more younger and fewer older employees, a result that is robust to controls for firm size, industry, geography, and year.

5. Evidence on mechanisms

In the previous section we show a positive relation between employee age and firm age. In this section, we seek to understand the drivers of this relation by exploring related empirical evidence. We first look at the relation between firm age and age of new hires. Second, we explore average wages for young and older workers at both young and older firms. Next, we look at the relation between firms’ potential for innovation and employee age. Finally, the correlation between firm outcomes and employee age is explored.

The intent of our analysis is not to identify one unique driver or even to provide a complete accounting of all the complex and jointly determined trends of young firms and young employees, a point emphasized by the fact that several of the identified mechanisms yield identical predictions. Instead, we document empirical evidence consistent with three non-mutually exclusive mechanisms. The results are consistent with the employee age-firm age pattern in the data for

\textsuperscript{15} We map NAICS codes to SIC codes for years 2002–2004.

\textsuperscript{16} We also estimate logistic regressions in which the dependent variable is the log odds function of the fraction of employees in a given age category. We find our results are very similar to the OLS regressions in which the fractions of employees of a given age are the dependent variables and report these estimates for easier discussion of economic magnitudes.
relation being driven, in part, by unique and valuable skills of young employees, greater risk tolerance of young employees, and the joint dynamics of young firms and young employees. Our analysis, however, does not allow us to definitively exclude alternative mechanisms that could also be at play, given the interrelatedness of various explanations of the employee age, firm age relation and the complex nature of employee-firm job matching.

5.1. Evidence from new hires

The joint dynamics mechanism argues the young employee, young firm relation is driven, at least in part, by the fact that young firms can only employ workers who have recently completed a job change and young employees change jobs more frequently. Thus, after limiting the sample to new hires who have all done a recent job search, we should observe a weaker relation between firm age and employee age. We explore this prediction in Table 3, Panel A using the full sample of all firms. In this table, the rows correspond to employee age categories (for new hires only) and the columns correspond to firm age categories.

First, we confirm the earlier assumption that younger workers make up a larger fraction of new hires. Column 6, which reports the age distribution of new hires averaged across all firm ages, shows that 50.5% of new hires are

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Fraction of employees aged 25–34</th>
<th>Fraction of employees aged 35–44</th>
<th>Fraction of employees aged 45–54</th>
<th>Fraction of employees aged ≥ 55</th>
</tr>
</thead>
<tbody>
<tr>
<td>Firm age 1–5 years</td>
<td>0.091*** (276.40)</td>
<td>0.040*** (107.34)</td>
<td>−0.053*** (−145.58)</td>
<td>−0.145*** (−352.65)</td>
</tr>
<tr>
<td>Firm age 6–10 years</td>
<td>0.060*** (195.50)</td>
<td>0.037*** (105.19)</td>
<td>−0.029*** (−82.86)</td>
<td>−0.114*** (−288.98)</td>
</tr>
<tr>
<td>Firm age 11–15 years</td>
<td>0.039*** (124.12)</td>
<td>0.023*** (65.12)</td>
<td>−0.008*** (−21.99)</td>
<td>−0.086*** (−212.10)</td>
</tr>
<tr>
<td>Firm age 16–20 years</td>
<td>0.019*** (66.83)</td>
<td>0.008*** (25.52)</td>
<td>0.005*** (14.28)</td>
<td>−0.049*** (−133.77)</td>
</tr>
<tr>
<td>Lagged log(1 + firm employees)</td>
<td>0.020*** (212.23)</td>
<td>0.005*** (46.01)</td>
<td>−0.011*** (−111.99)</td>
<td>−0.024*** (−222.06)</td>
</tr>
<tr>
<td>Industry fixed effects? Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>State fixed effects? Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year fixed effects? Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>N 16,343,058</td>
<td>16,343,058</td>
<td>16,343,058</td>
<td>16,343,058</td>
<td></td>
</tr>
<tr>
<td>R² 0.050</td>
<td>0.015</td>
<td>0.029</td>
<td>0.088</td>
<td></td>
</tr>
</tbody>
</table>

Panel B: Public firms only

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Fraction of employees aged 25–34</th>
<th>Fraction of employees aged 35–44</th>
<th>Fraction of employees aged 45–54</th>
<th>Fraction of employees aged ≥ 55</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time from IPO 1–5 years</td>
<td>0.098*** (19.37)</td>
<td>−0.020*** (−4.45)</td>
<td>−0.072*** (−18.33)</td>
<td>−0.069*** (−20.51)</td>
</tr>
<tr>
<td>Time from IPO 6–10 years</td>
<td>0.082*** (25.14)</td>
<td>−0.008*** (−2.69)</td>
<td>−0.065*** (−23.82)</td>
<td>−0.057*** (−22.28)</td>
</tr>
<tr>
<td>Time from IPO 11–15 years</td>
<td>0.005*** (19.26)</td>
<td>−0.002 (−2.28)</td>
<td>−0.047*** (−12.95)</td>
<td>−0.042*** (−18.35)</td>
</tr>
<tr>
<td>Time from IPO 16–20 years</td>
<td>0.028*** (13.29)</td>
<td>0.005*** (2.47)</td>
<td>−0.025*** (−12.70)</td>
<td>−0.022*** (−18.25)</td>
</tr>
<tr>
<td>Lagged log(1 + firm employees)</td>
<td>0.003*** (4.90)</td>
<td>−0.004*** (−6.19)</td>
<td>−0.005*** (−7.41)</td>
<td>−0.006*** (−10.00)</td>
</tr>
<tr>
<td>Industry fixed effects? Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>State fixed effects? Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year fixed effects? Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>N 37,359</td>
<td>37,359</td>
<td>37,359</td>
<td>37,359</td>
<td></td>
</tr>
<tr>
<td>R² 0.205</td>
<td>0.075</td>
<td>0.145</td>
<td>0.178</td>
<td></td>
</tr>
</tbody>
</table>
between 25 and 44 years of age, consistent with past findings. Second, we find a positive relation between age of new hires and firm age. Column 1 reports that 28.7% of new hires at firms aged one to five years are in the 25–34-year old category. Moving across the columns by firm age, we see a decline in the percentage of new hires aged 25–34. The percentage hits 25.3 when firms are aged over 20 years, as reported in column 5. We see a similar pattern for new hires aged 35–44, but with a smaller decline in percentage hired as firms age. The percentage of new hires aged 45–54 is relatively flat across firm age and the percentage of new hires 55 or older more or less increases by firm age.

Table 3, Panel B reports the relation between years since IPO and age of new hires for the sample of public firms. For firms that have gone public one to five years ago, 37.4% of new hires are aged 25–34. For public firms that have gone public over 20 years ago, only 30% of new hires are aged 25–34. In contrast, for public firms that have gone public one to five years ago, 13.3% of new hires are aged 45–54. For firms which have been public over 20 years ago, this percentage rises to 15.5%. Thus, the new hire employee-age firm-age relation is also present in the sample of larger public firms.

Table 4, Panel A shows that after controlling for firm size, industry, geography, and time, the relation between new hire age and firm age is robust. Column 1 shows that the percentage of new hires aged 25–34 in firms aged one to five years is three percentage points higher than in firms over 20 years in the full sample of privately and publicly held firms. Table 4, Panel B shows the results are similar when looking just at public firms. Column 1 shows that the percentage of new hires aged 25–34 in firms that have gone public one to five years ago is 4.4 percentage points higher than in firms that have gone public over 20 years ago. While we find a similar pattern between employee age and firm age when looking at new hires, the relation is less pronounced as compared to the sample of all employees. The more subtle relation in the sample of new hires is consistent with the joint dynamics mechanism. Moreover, the fact that we continue to observe a positive relation between firm age and new hire age is indicative that other mechanisms also play a role in explaining the firm age-employee age relation.

Looking at just new hires provides insight into the mechanisms behind the young employee, young firm relation, however, we are only able to observe the age distribution of new hires at the sample of firms which are hiring. New hires will include new employees to fill new jobs (expansion hires) as well as new employees to fill jobs that were vacated by departing employees (replacement hires). Since growing firms are hiring, by definition, this sample may be biased towards growing firms. On the other hand, Brown and Matsa (2013) find evidence...
indicating firms in distress experience higher employee turnover, suggesting declining firms may have large hiring needs as workers voluntarily leave to seek out better opportunities.

If the sample of firms with new hires has an over-representative distribution of either growing or non-growing firms, this could bias our results if the relation between firm age and employee age depends on firm growth rates. However, this does not appear to be the case. As reported in the Internet Appendix Tables IA.1 and IA.2, we find similar correlations between firm age and new hire age in the sample of growing and non-growing firms—suggesting differences in growth rates for the set of hiring firms will not bias the results.17

5.2. Evidence from wages

If young firms hire more young employees because young employees possess skills that are relatively more valuable to young firms, then young employees in young

---

17 There is a more complete discussion of the differences between the full sample and sample of firms with new hires in Section 6.
firms should receive greater compensation relative to young employees elsewhere. Moreover, young employees in young firms should earn compensation which is relatively more equal to their older colleagues, as compared to young employees at older firms. Alternatively, if positive assortative matching is driving the young employee, young firm relation, then we should observe relatively lower wages for young workers at young firms.

We explore the relation between firm age, employee age, and wages in Table 5. OLS estimates of the natural logarithm of the average wage per employee are reported for employees of all ages in column 1 and for employees in a specific age category in columns 2–5. The independent variables in the regressions include the four youngest firm age categorical variables as well as the lagged natural logarithm of one plus total firm employees, and industry, and state and year fixed effects. To preserve space, we only report the coefficient on the firms age one to five categorical variable which should be interpreted as the mean difference in the dependent variable for the youngest firm age category as compared to firms aged over 20 years, after controlling for firm size and industry, state and year fixed effects. Panels A and B explore wages for all firm employees. Panels C and D explore wages for new hires. Panels A and C use a sample of all public and private firms. Panels B and D use a sample of just publicly traded firms.

We see in Table 5, Panel A, column 1 that the average wage per employee, averaged across all employee age groups, is lower at firms aged one to five. The average employee working for a firm aged one to five years old earns 6.2% less than the average employee working at a firm aged over 20 years. In unreported results, there is a steady increase in wages as firms age.

Moreover, when we decompose average wages by employee age, we see that the picture is more complicated. For younger employees, i.e., those under age 45, the average wage per employee is higher in young firms, as reported in columns 2 and 3. On average, employees aged 25–34 earn 3.1% more and employees aged 35–44 earn 2.0% more at firms aged one to five years, relative to similarly aged employees in firms over 20 years old. We see a very different pattern for employees aged 45 and older in columns 4 and 5. Employees aged 45–54 are paid almost 10% less and employees aged 55 or older are paid nearly 24% less in firms aged one to five years relative to firms older than 20 years. These large negative wage premia for older employees in younger firms contribute to the average effect observed across all employees in column 1.

For public firms, as reported in Table 5, Panel B, the average employee at a firm which had its IPO in the last one to five years earns a higher wage, as compared to the average employee at a firm which had its IPO over 20 years ago. Furthermore, employees in the three youngest age groups all earn higher wages at the firms with the most recent IPOs as compared to similarly aged workers employed at the oldest firms. Employees aged 55 or older receive higher wages at firms which have been public for over 20 years vis-à-vis firms with recent IPOs.

These results are consistent with predictions of the skills mechanism. In contrast, the results do not directly support an argument that the firm age, employee age relation is attributed to a positive assortative matching mechanism.

While these wage results are consistent with the skills mechanism, they might also be explained by a tenure wage model. Lazear (1981) and Akerlof and Katz (1989) argue a firm may motivate young workers to work harder today at low wages if given the promise of higher future wages. Such implicit contracts will be more meaningful at firms where survival probabilities are higher. Given older firms have higher survival probabilities, a tenure wage model will predict relatively lower wages for young employees and relatively higher wages for older employees at older firms.

While we cannot completely exclude this alternative interpretation of the wage results, we document a number of results indicating a tenure wage model is not exclusively driving our results. For one, the tenure wage model predicts we should find a weaker pattern at public firms, given the difference in failure rates between young and old public firms is more modest relative to the comparable difference for private firms. Instead, we find an even larger wage premium for younger employees at young public firms.

Second, a wage tenure model predicts the relation would be weaker when exploring wages of new hires only. Looking first at new hires using the sample of all firms in Table 5, Panel C, we report that young new hires at young firms are paid relatively higher wages as compared to young new hires at firms over 20 years of age, consistent with the skills mechanism and a wage tenure argument. However, we also observe that older new hires receive higher wages at younger firms. This second finding is not consistent with either the skills or a wage tenure argument. The second finding is reduced when we exclude first-year firms, which suggests that older workers who start new firms pay themselves well in the first year. Older workers who are subsequently hired at young but not first-year firms are not as highly compensated.

Moreover, for public firms where we expect to observe few if any of these nascent firms, we find more similar results between wages for all employees and just new hires.

18 Becker (1962), Lazear (1981), and Akerlof and Katz (1989) document that at firms with a higher survival probability, younger employees are paid relatively less with a promise of higher payment if they are promoted within the firm.
19 Brown and Medoff (2003) find that the change in wages (from starting wage to date of survey) is higher per year of tenure at older firms, using a sample of approximately 1,000 workers from the Survey of Consumers. They argue these results support a tenure wage model. These results are also consistent with our skills mechanism. Younger workers at younger firms will be hired at a higher starting wage given their relatively greater value to young firms.
20 Bhattacharya, Borisov, and Yu (forthcoming) show that mortality rates of public firms hit a peak at 6% for three-year-old firms, then decline to less than 2% for firms 20 years or older. Haltiwanger, Jarmin, and Miranda (2012) show that the exit rates of private firms hit a peak for firms one to two years old at 15%, then decline to 2% for firms 16 years or older.
21 New hires of older workers at established firms should not receive a wage premium (relative to new hires of older workers at young firms) given these employees have not built up tenure.
Table 5

Wages by employee age and firm age.

Each column shows results from an OLS regression estimated using the log wage (in year 2005 dollars) per worker in a given category as the dependent variable. Column 1 includes all employees. Columns 2–5 restrict the sample to employees in a given age group. The independent variables are firm age categorical variables (i.e., firm age 1–5 years, firm age 6–10 years, firm age 11–15 years and firm age 16–20 years) and lagged natural logarithm of 1 plus total firm employees. The firm age category of over 20 years is omitted. Included in each specification are four-digit SIC code fixed effects, state fixed effects and year fixed effects. Only the coefficient on firm age 1–5 years and Time from IPO 1–5 years are reported to conserve space. The unit of observation is a firm-year. The data are taken from the union of the LEHD and LBD data sets between years 1992 and 2004. Panel A reports results for all employees using the full sample of private and public firms. Panel B reports results for all employees in the sample of publicly held firms only. Panel C reports results for the wages of new hires only using the full sample of private and public firms. Panel D reports results for the wages of new hires only using the sample of public firms only. t-Statistics adjusted for clustering by firm are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: All firms: Wages of all employees

<table>
<thead>
<tr>
<th>Firm age 1–5 years</th>
<th>Log(wage/employee) all employees</th>
<th>Log(wage/employee) aged 25–34</th>
<th>Log(wage/employee) aged 35–44</th>
<th>Log(wage/employee) aged 45–54</th>
<th>Log(wage/employee) aged ≥ 55</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>16,336,715</td>
<td>9,565,061</td>
<td>10,615,733</td>
<td>9,509,330</td>
<td>7,881,437</td>
</tr>
<tr>
<td>R²</td>
<td>0.280</td>
<td>0.213</td>
<td>0.227</td>
<td>0.234</td>
<td>0.230</td>
</tr>
</tbody>
</table>

Panel B: Public firms: Wages of all employees

<table>
<thead>
<tr>
<th>Time from IPO 1–5 years</th>
<th>Log(wage/employee) all employees</th>
<th>Log(wage/employee) aged 25–34</th>
<th>Log(wage/employee) aged 35–44</th>
<th>Log(wage/employee) aged 45–54</th>
<th>Log(wage/employee) aged ≥ 55</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>37,359</td>
<td>36,402</td>
<td>36,781</td>
<td>36,526</td>
<td>35,141</td>
</tr>
<tr>
<td>R²</td>
<td>0.206</td>
<td>0.208</td>
<td>0.217</td>
<td>0.211</td>
<td>0.176</td>
</tr>
</tbody>
</table>

Panel C: All firms: Wages of new hires

<table>
<thead>
<tr>
<th>Firm age 1–5 years</th>
<th>Log(wage/new hire) all employees</th>
<th>Log(wage/new hire) aged 25–34</th>
<th>Log(wage/new hire) aged 35–44</th>
<th>Log(wage/new hire) aged 45–54</th>
<th>Log(wage/new hire) aged ≥ 55</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>11,355,225</td>
<td>6,269,452</td>
<td>5,664,829</td>
<td>4,104,194</td>
<td>2,760,070</td>
</tr>
<tr>
<td>R²</td>
<td>0.209</td>
<td>0.199</td>
<td>0.203</td>
<td>0.193</td>
<td>0.171</td>
</tr>
</tbody>
</table>

Panel D: Public firms: Wages of new hires

<table>
<thead>
<tr>
<th>Time from IPO 1–5 years</th>
<th>Log(wage/new hire) all employees</th>
<th>Log(wage/new hire) aged 25–34</th>
<th>Log(wage/new hire) aged 35–44</th>
<th>Log(wage/new hire) aged 45–54</th>
<th>Log(wage/new hire) aged ≥ 55</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>36,384</td>
<td>34,863</td>
<td>34,679</td>
<td>33,161</td>
<td>28,655</td>
</tr>
<tr>
<td>R²</td>
<td>0.147</td>
<td>0.147</td>
<td>0.142</td>
<td>0.122</td>
<td>0.098</td>
</tr>
</tbody>
</table>

hires, consistent with the skills mechanism. In Table 5, Panel D, we report that younger new hires are paid relatively more and new hires aged 55 or older are paid relatively less at firms which had an IPO within the last five years vis-a-vis firms which had an IPO more than 20 years ago. The age groups that command the highest wage premia in young public firms are new hires aged 25–34 years, who earn almost 10% more in young public firms, as compared to firms that went public over 20 years ago, and new hires aged 35–44 years, who earn 11.6% more.

5.3. Evidence from firms’ potential for innovation

If younger workers join young firms because their skills are especially important to innovation, then the skills mechanism predicts we should find young firms which are more innovative employ a relatively greater share of younger workers. Proxying for innovation in young firms is difficult as innovation often takes years to be observed. Thus, instead of measuring innovation directly, we use a proxy for future expected innovation—the attainment of
VC financing. VC financing has been shown to be associated with rapid firm growth and innovation in the product market (e.g., Hellmann and Puri, 2000; Gompers and Lerner, 2001; Puri and Zarutskie, 2012).

In Table 6, we explore the relation between receiving VC financing and employee age at young firms. Columns 1 and 4 report marginal effects from probit models which estimate the probability a firm receives VC financing as a function of its initial characteristics. Fraction of employees aged 55 or older is the excluded group. The probit models are cross-sectional, and the number of observations is equal to the number of firms represented in each model. Column 1 estimates the probit model on all new firms in all industries. Given VC investment is concentrated in the “high-tech” industries of Computers, Electronics, Biotech, and Telecom, column 4 estimates the probit model only on these industries. We find that having more employees aged 25–34 compared to the excluded group of workers aged 55 or older is associated with a higher chance of receiving VC financing in the future.23 These results hold when considering all firms and just high-tech industries. Having more workers in the age category of 35–44 as compared to workers over 55 is also positively associated with VC financing in both samples.

Assuming VC-financed firms need more cutting-edge skills to innovate, as compared to non-innovative firms, then the results in Table 6 are consistent with the skills mechanism. However, we also acknowledge that firms which receive VC financing differ based on observable and unobservable characteristics and, thus, we cannot definitively exclude the possibility that these results are driven by an omitted variable correlated with VC financing and employee age. In the next section, we further test the relation between firm outcomes and employee age.

5.4. Evidence from firm outcomes

If young workers are more risk tolerant, as in Vroom and Pahl (1971) and Hensley (1977), then we expect that young employees match to the firms that present greater income or employment risk, such as those firms that display greater failure and growth rates. In Jovanovic (1982), firms learn that they are either efficient, in which case they grow, or that they are inefficient, in which case they exit. If young employees, who are also learning about their own skills, are relatively more likely to join firms that are learning about their abilities, then we should observe more young employees at firms with both higher growth rates and higher failure rates.

The skills mechanism suggests the same prediction regarding growth rates but the opposite prediction for failure rates. If young employees provide skills especially valuable to young firms, then we should observe higher growth and fewer failures among young firms with more young workers. Furthermore, this should be especially true in industries where human capital is advancing most quickly, suggesting relatively greater value to more recently acquired skills. The assortative matching mechanism predicts more failures and lower growth at firms with more young employees.

In columns 2 and 5 of Table 6, we report estimates from probit models which predict whether a new firms fails within five years as a function of the age distribution of its initial labor force, its initial size, and industry, state, and year fixed effects. Marginal effects are reported and the fraction of employees over 54 is the excluded group.24 In column 2, which estimates the model on all industries, we see that new firms with more workers aged 25–34, as compared to workers over 54, are more likely to fail. Alternatively, new firms with more workers 35–44, as compared to workers over 54, are less likely to fail. In column 5, which estimates the model only on firms in high-tech industries, we see that firms with more employees aged 25–34 and 35–44 are less likely to fail relative to firms with more workers aged 55 or older. For all industries, column 2, and high-tech industries, column 5, there is a strong relation between relatively more workers aged 45–54, as compared to workers aged 55 or older, and fewer failures. These results are consistent with the influence of multiple non-mutually exclusive mechanisms with competing predictions.

In columns 3 and 6 of Table 6, we regress the log of a firm’s employment growth rate over the first five years as a function of the firm’s initial labor force age distribution, total employment size, and state, industry, and year fixed effects. The number of observations, or firms, drops in these regressions because firms must survive at least five years to be included in the analysis. We see that new firms that survive for five years and employ relatively more workers aged 25–34 and 35–44, as compared to workers aged 55 or older, grow more quickly. These results hold when looking at the full sample of firms in column 3 and when limiting the sample to high-tech firms in column 6. These results support the skills, risk, and joint dynamics mechanisms.

6. Sample selection

Throughout the analysis, we make use of different sub-samples. We limit the sample to just public firms (e.g., Table 1, Panel B), firms which are hiring new workers (e.g.,

23 Because the probability of receiving VC is very low, less than 0.1% in the full sample of firms and only 1–5% in the sample of high-tech industries, the magnitudes of the marginal effects are small in columns 1 and 4.

24 The lower number of observations in column 2 compared to column 1 and in column 5 compared to column 4 of Table 6 reflect the fact that we drop firms for which we cannot observe the five-year failure rate due to right-hand censoring of the data. We find similar results when we account for censoring using a hazard model estimation.
Table 6
Initial employee age and new firm outcomes.

Columns 1 and 4 report probit estimates which predict whether a firm ever receives VC financing. Columns 2 and 5 report probit estimates which predict whether a firm fails within five years of being started. Marginal probabilities are reported followed by z-statistics in parentheses for these probit models. Columns 3 and 6 report OLS regression estimates in which the log 5-year employment growth rate, for surviving firms, is the dependent variable. The sample in column 3 is further limited to include only firms which survive for five years. Columns 4 and 6 mirror the regressions in columns 1–3 after limiting the sample to just firms in high-tech industries. High-tech is defined to include Biotech, Telecom, Computers, and Electronics. A firm is in the “Biotech” industry if its primary SIC code is 2830–2839, 3826, 3841–3851, 5047, 5048, 5122, 6324, 7352, 8000–8099, or 8730–8739 excluding 8732. A firm is in the “Telecom” industry if its primary SIC code is 3660–3669 or 4810–4899. A firm is in the “Computers” industry if its primary SIC code is 3570–3579, 5044, 5045, 5734, or 7370–7379. A firm is in the “Electronics” industry if its primary SIC code is 3600–3629, 3643, 3644, 3670–3699, 3825, 5065, or 5063.

### All industries

<table>
<thead>
<tr>
<th></th>
<th>(1) Receives VC</th>
<th>(2) Fails within 5 years</th>
<th>(3) 5-year employment growth rate</th>
<th>(4) Receives VC</th>
<th>(5) Fails within 5 years</th>
<th>(6) 5-year employment growth rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fraction employees aged 25–34 in firm’s first year</td>
<td>0.00074***</td>
<td>0.014***</td>
<td>0.285***</td>
<td>0.0018***</td>
<td>-0.024***</td>
<td>0.326***</td>
</tr>
<tr>
<td></td>
<td>(10.76)</td>
<td>(5.50)</td>
<td>(42.52)</td>
<td>(4.99)</td>
<td>(-2.83)</td>
<td>(12.56)</td>
</tr>
<tr>
<td>Fraction employees aged 35–44 in firm’s first year</td>
<td>0.00068***</td>
<td>-0.026***</td>
<td>0.166***</td>
<td>0.0018***</td>
<td>-0.050***</td>
<td>0.159***</td>
</tr>
<tr>
<td></td>
<td>(9.15)</td>
<td>(-9.91)</td>
<td>(25.11)</td>
<td>(4.91)</td>
<td>(-6.97)</td>
<td>(6.17)</td>
</tr>
<tr>
<td>Fraction employees aged 45–54 in firm’s first year</td>
<td>0.00056***</td>
<td>-0.032***</td>
<td>0.085***</td>
<td>0.0014***</td>
<td>-0.043***</td>
<td>0.101***</td>
</tr>
<tr>
<td></td>
<td>(7.65)</td>
<td>(-11.11)</td>
<td>(12.03)</td>
<td>(3.59)</td>
<td>(-4.66)</td>
<td>(3.61)</td>
</tr>
<tr>
<td>Log(1+firm employees) in firm’s first year</td>
<td>0.00039***</td>
<td>-0.067***</td>
<td>-0.222***</td>
<td>0.0017***</td>
<td>-0.035***</td>
<td>0.172***</td>
</tr>
<tr>
<td></td>
<td>(63.63)</td>
<td>(-78.92)</td>
<td>(-125.77)</td>
<td>(41.69)</td>
<td>(-14.12)</td>
<td>(-26.06)</td>
</tr>
<tr>
<td>Industry fixed effects?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>State fixed effects?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year fixed effects?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>N</td>
<td>1,048,366</td>
<td>814,074</td>
<td>301,909</td>
<td>158,643</td>
<td>91,323</td>
<td>35,857</td>
</tr>
<tr>
<td>Pseudo-R²/R²</td>
<td>0.355</td>
<td>0.032</td>
<td>0.353</td>
<td>0.337</td>
<td>0.052</td>
<td>0.040</td>
</tr>
</tbody>
</table>

Table 3, the cross-section of new firms (columns 1 and 4 of Table 6), new firms which survive for five years (columns 3 and 6 of Table 6), and firms just in high-tech industries (columns 4–6 of Table 6). Repeating similar tests across various samples, such as comparing the firm age-employee age relation at all firms, then at the sample of public firms, allows us to explore the robustness of the results. Focusing on subsamples allows us to gain insight into the importance of specific mechanisms. However, any interpretation is complicated given the subsamples may differ along dimensions other than those highlighted in the tests. In this section, we discuss these concerns in more detail.

6.1. Public firms

Most young firms are private. It is a unique feature of Census data and, subsequently of this analysis, that private firms are included. Since private firms dominate the full sample, we also repeat our analyses on the subset of public firms. Many research samples are limited to using public firms and this allows our results to be more directly comparable to these studies.

In Table A3, Panel A, we report summary statistics for the sample of all firms. In Table A4, we report summary statistics for the sample of public firms. Public firms are larger, as measured by employee count, and pay higher wages, as compared to the full sample. The pattern is present in each firm age category. With public firms, we can also document additional firm characteristics. We report increasing firm size and leverage and declining Tobin’s Q with firm age. Suffice to say, the sample of all firms, dominated by small private firms, and the sample of public firms are different on multiple dimensions. We argue that it is, therefore, quite striking that we find similar patterns between firm age and employee age across the two samples.

6.2. Firms which are hiring

In Table A3, Panel B, we report summary statistics for the sample of firms which are hiring. As discussed in Section 5.1, the sample of firms with new hires may differ from the full sample as these firms may be associated with different growth rates.25

---

25 This is true only for the sample of private firms. Of the sample of public firms, 97% hire in a given firm-year, suggesting very little difference between the samples of public firms used in tests with all employees and tests with new hires.
Employee growth is defined as the annual change in firm-level employees divided by the average employment at the firm over the two years during which the change is measured, as in Haltiwanger, Jarmin, and Miranda (2012). To benchmark our employment growth results, we compare them to Haltiwanger, Jarmin, and Miranda (2012), however, a few differences between the samples complicate the comparison. For one, Haltiwanger et al. (2012) drop first-year firms from their analysis of growth and instead look at growth rates for firms two years or older.26 Furthermore, Haltiwanger, Jarmin, and Miranda (2012) use employee-weighted averages. We report establishment-weighted averages.

Similar to Haltiwanger et al. (2012), we observe higher employment growth in firms under age five. After age five, there appears to be little to no relation between firm age and mean employment growth. We find a slightly lower mean employment growth rate, vis-a-vis Haltiwanger, Jarmin, and Miranda (2012). The difference is likely due to the different averaging methodology. Haltiwanger, et al. (2012) note that, after controlling for firm age, firm size is positively correlated with employment growth.

We find a lower growth rate in Table A3, Panel B when we limit the sample to hiring firms. This lower growth rate is likely due to a combination of some new hires constituting replacement hires and the nature of the composition of the data set. The difference in growth rates would be of concern to our interpretation if there are different patterns associated with replacement hires as compared to expansion hires. We explore this issue by repeating key tables in the paper and limiting the sample to either just growing firms or just non-growing firms. These results are reported in Tables IA.1 and IA.2 of the Internet Appendix. We find generally similar patterns in the sample of just growing and just non-growing firms. In particular, if we repeat the regressions in Table 4, Panel A using only growing firms, the coefficient on the age 1–5 dummy is 0.034*** for new hires aged 25–34, 0.013*** for new hires aged 35–44, –0.007*** for new hires aged 45–54, and –0.023*** for new hires aged ≥55. If we repeat the regressions in Table 4, Panel A using only non-growing firms, the coefficient the on age 1–5 dummy is 0.032*** for new hires aged 25–34, 0.001*** for new hires aged 35–44, –0.022*** for new hires aged 45–54, and –0.038*** for new hires aged ≥55.

6.3. Firms for which we observe wages

Additional sample restrictions apply when we consider wages. As long as a firm employs (hires) at least one worker, we can estimate the fraction of employees (new hires) in each age group. However, we need the firm to employ (have hired) at least one worker in a given age group to be able to estimate the average wages for employees (new hires) in that age group. As such, the relation between firm age and wages by employee age involves different samples. Given the evidence in Table 6 that firms which employ more young workers are different, the wage results we report in Table 5 could be biased.

We address this issue in Table 7 by exploring the relative wage premium between young and old workers employed at the same firm. We estimate the within-firm difference in average wages per employee (new hires) between older workers (those aged 45–54 and 55 or older) and younger workers (those aged 25–34) and test whether this within-firm wage differential can be explained by firm age. Table 7, Panel A presents estimates for the full sample of private and public firms; Panel B presents estimates for the subsample of public firms. The regressions include industry, state and year fixed effects, and firms aged over 20 years is the omitted dummy variable.

In the full sample, we see that the wage differential between employees aged 45–54 and employees aged 25–34 is 15.7% lower for firms aged 1–5, as compared to firms aged over 20 years, as reported in column 1. Moreover, there is an increase in the spread between these two groups of employees as firms age. The pattern is even more dramatic if we look at the wage differential between employees aged 55 or older and employees aged 25–34 in column 2. When we focus on new hires in column 3, we actually see that the wage gap between 45 and 54 year-old new hires and 25–34 year-old new hires widens in the youngest firms. However, this is driven by first-year firms where older workers who begin their employment in the year of the firm's birth receive high wages. When we drop first-year firms from the sample and only include firms aged two to five years, we see that the coefficient on the firm age one to five dummy becomes negative and significant at the 1% level. The wage spread between new hires 55 or older and new hires aged 25–34 also becomes significantly negative when we drop first-year firms.

Focusing on the estimates for the subsample of public firms in Table 7, Panel B, we see that the spread in wages for employees aged 45–54 and employees aged 25–34 narrows by 5% for the firms which recently went public as compared to firms which went public over 20 years ago, both when we examine wages of all employees and of new hires only. The wage spread between workers aged 55 or older and workers aged 25–34 narrows by 20.9% and 11.6% for all employees and new hires, respectively, for the youngest vis-a-vis the oldest public firms.

These results indicate that the positive wage spread documented by labor economists between older and younger employees (e.g., Ben-Porath, 1967; Murphy and Welch, 1990, 1992) is narrower within young firms, with young employees earning relatively more equal pay with older employees in the same firm. This is consistent with the relative skill/productivity gap between younger and older workers being smaller at young firms as compared to older firms. Moreover, the results are consistent with the earlier results, suggesting sample differences are not driving our results. However, we acknowledge that while these tests control for one bias (firms with different characteristics may hire different age workers), it introduces a different bias (to be in the sample, the firm must be hiring both old and young workers).

6.4. Firm fixed effects

We primarily rely on cross-sectional variation for identification throughout our tests. While our data set
Table 7
Within-firm differences in wages of younger and older employees.

OLS regressions are estimated using the difference in log wage (in year 2005 dollars) per employee between two employee age categories as the dependent variable. The independent variables are firm age categorical variables and lagged natural logarithm of 1 plus total firm employees. The firm age category of over 20 years is omitted. Included in each specification are four-digit SIC code fixed effects, state fixed effects, and year fixed effects. The unit of observation is a firm-year. t-Statistics adjusted for clustering by firm are reported in parentheses. Panel A reports results for all firms in the union of the LEHD and LBD databases between years 1992 and 2004. Panel B reports results for the sample of public firms. The first two columns report regressions using wages for all employees. The second two columns report regressions using wages for new hires only. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: All firms

<table>
<thead>
<tr>
<th>Firm age 1–5 years</th>
<th>(1) Difference in log(wage) employee</th>
<th>(2) Difference in log(wage) employee</th>
<th>(3) Difference in log(wage/new hire) employee</th>
<th>(4) Difference in log(wage/new hire) employee</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>between 45–54 and 25–34 employee age categories</td>
<td>between 55 and 25–34 employee age categories</td>
<td>between 45–54 and 25–34 new hire age categories</td>
<td>between 55 and 25–34 new hire age categories</td>
</tr>
<tr>
<td>Firm age 6–10 years</td>
<td>-0.157*** (93.66)</td>
<td>-0.314*** (148.66)</td>
<td>0.006*** (3.47)</td>
<td>-0.017*** (3.75)</td>
</tr>
<tr>
<td>Firm age 11–15 years</td>
<td>-0.097*** (62.51)</td>
<td>-0.263*** (134.11)</td>
<td>-0.019*** (11.82)</td>
<td>-0.017*** (7.65)</td>
</tr>
<tr>
<td>Firm age 16–20 years</td>
<td>-0.021*** (13.38)</td>
<td>-0.182*** (90.55)</td>
<td>-0.014*** (7.98)</td>
<td>-0.012*** (4.97)</td>
</tr>
<tr>
<td>Lagged log(1 + firm employees)</td>
<td>-0.024*** (16.17)</td>
<td>-0.080*** (43.88)</td>
<td>-0.009*** (5.19)</td>
<td>-0.007*** (2.94)</td>
</tr>
<tr>
<td></td>
<td>(34.03)</td>
<td>(32.58)</td>
<td>(19.65)</td>
<td>(10.49)</td>
</tr>
</tbody>
</table>

Industry fixed effects? Yes Yes Yes Yes
State fixed effects? Yes Yes Yes Yes
Year fixed effects? Yes Yes Yes Yes
N 6,139,985 5,076,116 2,573,540 1,765,011
R² 0.026 0.045 0.008 0.029

Panel B: Public firms

<table>
<thead>
<tr>
<th>Time from IPO 1–5 years</th>
<th>(1) Difference in log(wage) employee</th>
<th>(2) Difference in log(wage) employee</th>
<th>(3) Difference in log(wage/new hire) employee</th>
<th>(4) Difference in log(wage/new hire) employee</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>between 45–54 and 25–34 employee age categories</td>
<td>between 55 and 25–34 employee age categories</td>
<td>between 45–54 and 25–34 new hire age categories</td>
<td>between 55 and 25–34 new hire age categories</td>
</tr>
<tr>
<td>Time from IPO 6–10 years</td>
<td>-0.049*** (2.95)</td>
<td>-0.209*** (9.34)</td>
<td>-0.050*** (3.40)</td>
<td>-0.116*** (5.19)</td>
</tr>
<tr>
<td>Time from IPO 11–15 years</td>
<td>0.003 (0.28)</td>
<td>-0.112*** (7.48)</td>
<td>-0.026*** (2.57)</td>
<td>-0.091*** (6.12)</td>
</tr>
<tr>
<td>Time from IPO 16–20 years</td>
<td>0.013 (1.26)</td>
<td>-0.054*** (3.99)</td>
<td>-0.026*** (2.76)</td>
<td>-0.064*** (4.64)</td>
</tr>
<tr>
<td>Lagged log(1 + firm employees)</td>
<td>-0.002 (0.26)</td>
<td>-0.022*** (2.22)</td>
<td>-0.003 (0.39)</td>
<td>-0.027*** (2.48)</td>
</tr>
<tr>
<td></td>
<td>(0.018*** (9.07)</td>
<td>-0.027*** (9.49)</td>
<td>-0.001 (0.50)</td>
<td>-0.004 (1.29)</td>
</tr>
</tbody>
</table>

Industry fixed effects? Yes Yes Yes Yes
State fixed effects? Yes Yes Yes Yes
Year fixed effects? Yes Yes Yes Yes
N 35,860 34,602 32,426 28,266
R² 0.082 0.075 0.029 0.033

Consists of millions of firms, it contains at most 13 years of data for a given firm. Perhaps young firms are not themselves unique but instead firms created in the 2000s (firms which will be included in the youngest age group) are different vis-a-vis firms created in the 1980s (firms which will be included in the oldest age groups) and these differences in firm characteristics over time are driving our results. To address this alternative interpretation of our results, we repeat key results after including firm fixed effects in the regressions. Including firm fixed effects involves a trade-off. On one hand, by including firm fixed effects, we dramatically reduce the power of our tests due to the short time series available. However, without firm fixed effects, we cannot exclude the possibility that our

27 The average firm is observed for 4.6 years. This is partially driven by the high attrition rates of firms but is also greatly determined by the limited availability of LEHD data.
results are driven by changes in the distribution of firm characteristics over time.

In Table IA.6 of the Internet Appendix, we repeat Table 2 with firm fixed effects. We still find a positive employee age-firm age pattern, albeit with weaker statistical significance. For example, we find that the share of workers aged 25–34 is 1.5 percentage points higher when firms are aged 1–5 years compared to when they are older than 20 years, when considering the full sample. This difference is 1.4 percentage points when estimated using the sample of public firms.

In Table IA.7 of the Internet Appendix, we repeat Table 4 with firm fixed effects. We find that the share of new hires aged 25–34 is 0.3 percentage point higher when firms are aged one to five years, compared to firms older than 20 years, in the full sample of firms. For the sample of public firms, the difference is 0.6%, however, this difference is not statistically significant. Overall, we find weaker but still often significant results when looking at new hires using the full sample. We find insignificant differences when looking at new hires and the full sample. The loss of power is likely driven by the limited within-firm variation in age in this sample.

We also find that the wage results reported in Table 5 Panel A are generally robust to the inclusion of firm fixed effects in the estimation. Results are reported in Table IA.8 of the Internet Appendix. For example, in the sample of all firms, employees between 25 and 34 years of age receive wages 10% higher when employed at a firm aged one to five as compared to a firm over 20 years, on average. Wage results using new hires or public firms are weaker once firm fixed effects are included.

7. Does the supply of young workers affect new firm creation?

In the previous sections, we showed that young firms disproportionately employ young workers and presented evidence that the unique skills of young employees, greater risk tolerance of young employees, and the joint dynamics of young firms and employees can partially explain this relation. If young employees are a critical component for young firms, then we should expect to find that exogenous changes to the supply of young workers affect new firm creation. In this section, we test this prediction.

We use the historic ratio of adolescents and young adults in the population as a proxy for the number of young workers in an area ten or 20 years later. Specifically, we calculate the ratio of the population in a state between 15 and 24 years of age and between five and 15 years of age, as compared to the population between 15 and 54 years of age, using data provided by the US Census of Population. We find the historic ratio, lagged ten years, is a strong predictor of age groups in the same state, after controlling for state fixed effects.28 We argue this lagged ratio not only reflects the supply of young workers in a given state, but most importantly, is unlikely to be driven by current job opportunities, at least when used in a regression after controlling for time-invariant differences across states.

The typical approach would be to use these variables as instruments in a two-stage least squares estimation. However, the Census does not provide state-level age distribution data from 1991 onwards. Thus, in order to maximize the sample years used in the estimation, we instead use the lagged age ratios directly as a proxy for the supply of young workers in a given state, which is not likely to be driven by current job opportunities.

Table 8 presents the OLS regressions which predict new firm creation rates in a state and state-industry pair as a function of the fraction of the population aged 15–24, lagged by ten years, or the fraction of the population aged 5–14, lagged by 20 years, as well as state and year fixed effects. The sample period in these regressions is 1980–2000. The first two specifications present estimates for new firm creation in all industries. We see that an increase in the fraction of the population that is young is associated with an increase in new firm creation rates. In particular, an increase in the fraction of the population aged 15–24 ten years ago by five percentage points is associated with an increase in the new firm creation rate of about half a percentage point. We find similar, but insignificant, results when we consider the fraction of the population aged five to 14 20 years ago.

The remaining specifications estimate the impact of the share of youth in the population on new firm creation rates in the high-tech industries of Biotech, Electronics, and Telecom. We see that, in the case of Biotech and Electronics, an increase in the share of youth in the population has a positive impact on new firm creation rates. For example, an increase in the share of youth in the population by 5% leads to an increase in the rate of new firm creation in the Electronics industries of between one and two percentage points. These results suggest that the supply of young workers does impact new firm creation, especially in industries in which unique skills of young workers are demanded by young firms.

The evidence presented in this section further supports the argument that young workers are a necessary ingredient for the creation and growth of new firms, due to their unique skills or willingness to work for new ventures, and buttresses our previous results that young, high-growth firms disproportionately hire younger employees and pay them higher relative wages.

8. Conclusion

We present large-scale evidence that young employees are an important component in the creation and growth of young firms. We first show that young firms disproportionately employ younger workers. Around 27% of employees in firms aged one to five years are between 25 and 34 years old, and over 70% are under the age of 45. In contrast, in established firms that have been in existence for over 20 years, fewer than 18% of employees are between the ages of 25 and 34, while nearly half are over the age of 45.

We argue that the positive relation between firm age and employee age is due, at least in part, to a combination
of unique skills and greater risk tolerance of young employees, and the joint dynamics of young firms and young workers. Consistent with young employees having unique skills valued more highly in young firms, we find that young employees in young firms command higher wages that are relatively more equal to older employees within the same firm. Further, we show that young employees disproportionately join young firms which have greater potential for innovation, as proxied by receiving VC financing. We present evidence that greater risk tolerance of young employees can partially explain the positive relation between firm age and employee age. We show that the joint dynamics of young firms and young workers can mean for new firm creation rates and the rate of new firm creation, especially in high-tech industries.

Finally, we find evidence that the supply of young workers affects the rate of new firm creation, with fewer new firms being created when there are fewer available young workers in a region. Using historical demographic information on the relative ratio of youth in a state as a predictor for the ratio of younger to older workers ten and 20 years later, we argue that a causal relationship exists between the supply of young workers and the rate of new firm creation, especially in high-tech industries.

Our findings also point to future research questions. How might the need to attract and compensate young employees, or employees who share similar attributes or skills, influence the organizational structure or compensation schemes used by young firms? How might it influence the financing choices of young firms or otherwise interact with financial constraints? Finally, what do shifting workforce demographics mean for new firm creation rates and their subsequent dynamics?

Appendix A. Matching establishments between the LEHD and the LBD and sample characteristics of the matched sample

In this appendix, we describe how we match establishments in the LEHD to establishments in the LBD in more detail. We also discuss differences in the matched LBD-LEHD sample as compared to the full LBD and potential biases in earlier results associated with using different matching criteria or due to the time-series properties of the LEHD data.

A.1. Matching establishments in the LBD and the LEHD

We link business establishments in the QWI to business establishments in the LBD using the Business Register Bridge. The Business Register Bridge matches business establishments across the two databases using federal EIN, industry, state, and county of the establishments. The Business Register provides details to match observations based on 15 combinations of EIN, industry, state, and county. We use five of these matches based on EIN, industry, state, and county as well as EIN, state, and county; EIN, and state; EIN and four-digit NAICS; and EIN only. See Abowd et al. (2006) for a more detailed description of the crosswalk files.
Table A1
Matching method for establishments in the LEHD and the LBD.

This table reports the percentage of establishments in the LEHD that are matched to the LBD by the matching method used. Reported in the first column are the percentages of matched establishments in the period 1992–2001, for which SIC codes are used to identify an establishment’s industry. Reported in the second column are the percentages of matched establishments in the period 2002–2004, for which NAICS codes are used to identify an establishment’s industry.

| Percentage of establishments matched on EIN, state and county, three-digit SIC code | 6% | 7% |
| Percentage of establishments matched on EIN, state and county, two-digit SIC code | 7% | 12% |
| Percentage of establishments matched on EIN, state and county, one-digit SIC code | 8% | 6% |
| Percentage of establishments matched on EIN and state and county | 13% | 17% |

Table A2
Match rates between establishments in the LEHD and LBD.

This table reports the percentage of establishments in the LBD that are matched to the LEHD by year, for those state-years included in the LEHD. Over the period 1992–2001, establishments in the LEHD are first matched to the LBD using EIN, state and county, and four-digit SIC code, then using EIN, state and county, and three-digit SIC code, then using EIN, state and county, and two-digit SIC code, then using EIN, state and county, and one-digit SIC code, and finally, using just EIN and state and county. Over the period 2002–2004, establishments in the LEHD are first matched to the LBD using EIN, state and county, and five-digit NAICS code, then using EIN, state and county, three-digit NAICS code, then using EIN, state and county, two-digit NAICS code, then using EIN, state and county, and one-digit NAICS code, and finally, using just EIN and state and county.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1992</td>
<td>67%</td>
<td>65%</td>
<td>67%</td>
<td>66%</td>
<td>67%</td>
<td>69%</td>
<td>66%</td>
<td>66%</td>
<td>65%</td>
<td>73%</td>
<td>49%</td>
<td>55%</td>
<td>55%</td>
<td>63%</td>
</tr>
</tbody>
</table>

We first attempt to match establishments using the more precise industry information, e.g., EIN, state and county, and four-digit SIC code or five-digit NAICS code, and then relax the industry precision until we match establishments only on EIN and state and county. If multiple establishments in the LEHD match to one or more establishments in the LBD, we aggregate information for the LEHD establishments and assign this information to the aggregate LBD establishments to which they match. We then aggregate establishment-level variables up to the firm level to create a panel data set consisting of 4,374,025 firms tracked over the period 1992–2004, for a total of 20,185,572 firm-year observations.

Table A1 lists the percentage of matched establishments in the LEHD by matching criteria of EIN, state and county, and industry. The largest share of matches takes place using the most detailed industry information. Over the period 1992–2001, 66% of establishments are matched using EIN, state and county, and four-digit SIC code. Over the period 2002–2004, 58% of establishments are matched using EIN, state and county, and five-digit NAICS code. Smaller percentages of establishments are matched based on less detailed industry information. The next largest share of matches are made using just EIN and state and county, with 13% of matches over the period 1992–2001 and 17% of the matches over the period 2002–2004 being made this way. We do not match based on coarser information than these combinations of EIN, state and county, and industry to ensure the relative accuracy of our matches.

Our matching procedure yields relatively high match rates. Table A2 presents the percentage of establishments in the LBD that are located in states and years covered by the LEHD, and hence that could potentially be matched to the LEHD, that we do in fact match to LEHD establishments. Our overall match rate across all years is 63%, as seen in the last column of Table A2. Match rates are higher over the period 1992–2001, ranging from 65% to 73%, compared to the period 2002–2004, in which match rates range from 49% to 55%. This likely reflects the addition of new establishments to the LBD in 2002 as well as less accuracy in industry and EIN matching amongst these new establishments. The match rates for establishments of public firms, i.e., those whose parent firms are matched to Compustat, have lower match rates, averaging 35%. This reflects the difficulty in matching establishments that are part of multi-unit firms, which may themselves have multiple EINs and operate in multiple locations and industry segments.

A.2. Comparing the matched LEHD-LBD data and the LBD

We next describe the characteristics of the matched LEHD-LBD data set and discuss possible sample selection biases. Table A3, Panel A presents means of several key variables for the sample of firms that are used in the estimation of the employee age regressions in Table 2, Panel A. Means are presented for all firms in the sample and also by firm age category.
We see that the average number of employees for a firm in the matched LEHD-LBD database is 26. The average firm has 22 employees in the broader LBD, as computed by the authors. This difference is statistically significant at the 1% level as determined by a difference in means t-test allowing for unequal sample variances. The average payroll per employee, reported in constant year 2005 dollars, for a firm in the LEHD-LBD panel of 27,249 is slightly lower than that in the full LBD of $31,000. This difference is statistically different at the 5% level as determined in means test allowing for unequal sample variances.

Although the differences in employment and wages between the matched LEHD-LBD sample and the broader LBD are statistically significant, they are of modest economic magnitude. The difference in firm size may be driven by the fact that the LEHD states tend to have larger firms vis-a-vis firms in non-LEHD states. These comparisons suggest that there are no major selection biases with regard to wages between our matched sample and the larger LBD.

Table A4 presents means for the subsample of public firms in the matched LEHD-LBD sample. The average number of employees reported in the first row of Table A4 reflects only the employees in the matched public establishments, and excludes those establishments in the same public firm that are not matched. The average, 2,938.5, is significantly smaller than the average number of employees at public firms over our sample period of around 10,000 employees.

When we examine firm characteristics taken from Compustat data, such as total assets and sales, however, we see that the public firms included in our sample are representative of the larger population of public firms over our sample period. Therefore, while we only observe employee characteristics in a fraction of the establishments in the public firms that we can match to the LEHD, Table A3
Summary statistics for firms included in the main estimation sample.

This table reports means for key variables for the samples of firms used in earlier analysis. Panel A reports means for the sample of 16,343,058 firm-year observations in Table 2, Panel A. Panel B reports means for the sample of 11,404,068 firm-year observations in Table 4, Panel A. Employee growth is defined as the annual change in firm-level employees divided by the average employment at the firm over the two years during which the change is measured. For first-year firms, employment in the first year is set equal to zero. For firms that exit the sample, employment in the second year is set equal to zero. Payroll per employee and payroll per new hire are reported in constant year 2005 dollars. Means are reported for all firms in the first column and then by firm age category in subsequent columns as denoted by the column headings.

<table>
<thead>
<tr>
<th>Firm age</th>
<th>All ages</th>
<th>Ages 1–5</th>
<th>Ages 6–10</th>
<th>Ages 11–15</th>
<th>Ages 16–20</th>
<th>Ages &gt; 20</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
</tr>
<tr>
<td>Number of employees</td>
<td>26.0</td>
<td>10.2</td>
<td>11.7</td>
<td>16.0</td>
<td>29.2</td>
<td>66.4</td>
</tr>
<tr>
<td>Employee growth (including firm deaths)</td>
<td>−0.030</td>
<td>0.230</td>
<td>−0.022</td>
<td>−0.031</td>
<td>−0.021</td>
<td>−0.040</td>
</tr>
<tr>
<td>Payroll per employee</td>
<td>27,249</td>
<td>24,903</td>
<td>25,920</td>
<td>27,799</td>
<td>30,115</td>
<td>29,518</td>
</tr>
</tbody>
</table>

Panel B: Hiring firms

| Number of employees | 38.6 | 12.5 | 17.5 | 24.4 | 45.2 | 104.4 |
| Payroll per new hire | −0.048 | 0.201 | −0.072 | −0.071 | −0.050 | −0.063 |

Table A4
Summary statistics for public firms only.

This table reports means for key variables for the sample of public firms included in the regressions reported in Table 2, Panel B. Means are reported for all firms in the first column and then by firm age category, defined in terms of years from initial public offering (IPO) in subsequent columns as denoted by the column headings. Payroll per employee and payroll per new hire are reported in constant year 2005 dollars. Total assets, sales, and market capitalization are reported in millions of constant year 2005 dollars. Tobin’s Q is defined as fiscal year-end market value of equity plus market value of preferred stock plus total liabilities divided by total assets. Leverage is defined as total debt divided by total assets.

<table>
<thead>
<tr>
<th>Years since IPO</th>
<th>All years (1)</th>
<th>1–5 Years (2)</th>
<th>6–10 Years (3)</th>
<th>11–15 Years (4)</th>
<th>16–20 Years (5)</th>
<th>&gt; 20 Years (6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of employees in matched LEHD-LBD establishments</td>
<td>2,938.5</td>
<td>461.9</td>
<td>618.6</td>
<td>863.8</td>
<td>2,526.8</td>
<td>5,010.2</td>
</tr>
<tr>
<td>Payroll per employee</td>
<td>54,101</td>
<td>69,467</td>
<td>63,488</td>
<td>57,447</td>
<td>53,563</td>
<td>49,445</td>
</tr>
<tr>
<td>Payroll per new hire</td>
<td>40,482</td>
<td>48,466</td>
<td>45,643</td>
<td>42,134</td>
<td>38,895</td>
<td>38,124</td>
</tr>
<tr>
<td>Total assets</td>
<td>5,671.5</td>
<td>882.0</td>
<td>1,026.9</td>
<td>2,140.6</td>
<td>5,481.7</td>
<td>9,208.8</td>
</tr>
<tr>
<td>Sales</td>
<td>2,571.4</td>
<td>410.5</td>
<td>506.9</td>
<td>1,004.6</td>
<td>2,495.3</td>
<td>4,141.5</td>
</tr>
<tr>
<td>Market capitalization</td>
<td>3,309.0</td>
<td>905.0</td>
<td>961.5</td>
<td>1,506.8</td>
<td>2,849.1</td>
<td>5,321.1</td>
</tr>
<tr>
<td>Tobin’s Q</td>
<td>1.67</td>
<td>2.76</td>
<td>2.35</td>
<td>2.03</td>
<td>1.51</td>
<td>1.27</td>
</tr>
<tr>
<td>Leverage</td>
<td>0.219</td>
<td>0.173</td>
<td>0.198</td>
<td>0.245</td>
<td>0.221</td>
<td>0.220</td>
</tr>
</tbody>
</table>
the public firms included in the matched sample are representative of the broader population of public firms.

A.3. Does the LBD-LEHD matching criteria used matter?

As reported in Table A1, more than half of our matched LBD-LEHD sample is created using the most stringent matching criteria. To confirm that including observations matched using less stringent matching criteria is not biasing our results, we repeat our employee age and wage regressions (Tables 2, 4, and 5) but only include firms that were matched using the most stringent matching criteria for all establishments, i.e., all establishments must be matched based on EIN, state and county, and four-digit SIC code or five-digit NAICS code. We find our results are robust. For example, if we repeat the regressions in Table 4, Panel A using only observations which were matched using the most stringent match, the coefficient on the age one to five dummy is 0.030*** for new hires aged 25–34, 0.007*** for new hires aged 35–44, –0.016*** for new hires aged 45–54, and –0.031*** for new hires aged $\geq 55$.30

A.4. Do the time-series properties of the LEHD matter?

A second possible concern with the LBD-LEHD matching is that states join the LEHD program at different times. As such, we have different time-series properties for establishments located in different states. To ensure that the timing of states joining the LEHD does not bias our results, we repeat our employee age and wage regressions (Tables 2, 4, and 5) but restrict our sample to be a balanced panel of state-years from 1997 to 2004. We again find our results are robust. For example, if we repeat the regressions in Table 4, Panel A using a sample with a balanced panel of state-years, the coefficient on the age 1–5 dummy is 0.033*** for new hires aged 25–34, 0.011*** for new hires aged 35–44, –0.013*** for new hires aged 45–54, and –0.027*** for new hires aged $\geq 55$.

References


Hensley, W., 1977. Probability, personality, age, and risk-taking. The Journal of Psychology 95, 139–145.

---

30 We chose not to clear these results to reduce the total number of samples submitted through the Census clearance process and due to the fact that the results were so similar to the estimates we report in the paper using all matched establishments.