High Growth Young Firms: Contribution to Job Growth, Revenue Growth and Productivity

December 2014

John Haltiwanger, Ron S Jarmin, Robert Kulick and Javier Miranda*

Recent research shows that the job creating prowess of small firms in the U.S. is better attributed to startups and young firms that are small. But most startups and young firms either fail or don't create jobs. A small proportion of young firms grow rapidly and they account for the long lasting contribution of startups to job growth. High growth firms are not well understood in terms of either theory or evidence. Although the evidence of their role in job creation is mounting, little is known about their life cycle dynamics, or their contribution to other key outcomes such as real revenue growth and productivity. In this paper, we explore these issues using U.S. Census Bureau Longitudinal Business Database combined with evidence from the Census Business Register to explore the role of high growth young firms further. In exploring these issues, we use new real revenue and productivity measures developed from the integration of these comprehensive databases tracking U.S. firms and establishments.

* University of Maryland and NBER, U.S. Census Bureau, University of Maryland, and U.S. Census Bureau, respectively. We thank David Brown for helpful comments. We thank the Kauffman Foundation for financial support. Any opinions and conclusions expressed herein are those of the author(s) and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed.

I. Introduction

Business startups and high growth young firms disproportionately contribute to job creation in the U.S. In a typical year, startups account for about 10 percent of firms and more than 20 percent of firm level gross job creation. Less well known is that most U.S. business startups exit within the first ten years, and the median surviving young businesses do not create jobs but remain small. But a small fraction of young firms create jobs rapidly and contribute substantially to job creation. These high growth young firms are the reason that startups make a long lasting contribution to net job creation.[1]

Most of the limited evidence on high growth firms has been about their contribution to job creation. While the latter is clearly important, less is known about the nature of their contribution to real revenue (i.e., output) and productivity growth. The reason is primarily data limitations. For the U.S., substantial progress has been made in developing longitudinal business databases that permit tracking growth and survival of businesses in terms of jobs. For studies of real revenue and productivity, much of the focus has been on the manufacturing sector with some limited analysis of the retail trade sector using Economic Census data available every five years. In the current paper, we develop and report new findings using data that tracks the real revenue growth at the firm level for the entire U.S. private sector on an annual basis. To our knowledge, this is the first database at the firm level that tracks both revenue and employment outcomes for all types of firms in the private sector on an annual basis.[2] This enables us to study the

[1] This discussion is based on Haltiwanger, Jarmin and Miranda (2013) and Decker, Haltiwanger, Jarmin and Miranda (2014). Note that the statistic that startups account for more than 20 percent of firm level job creation is based on gross job creation by firms, not establishments. Startups account for slightly less than 20 percent of establishment level job creation.

[2] For publicly traded firms, COMPUSTAT provides a rich source of revenue, asset and other data. The quinquennial economic censuses can be used to provide revenue data for most sectors every five years. Annual surveys of specific sectors can be used to generate samples of firms for most sectors but they are not intended for longitudinal analysis at the firm level.

contribution of young high growth firms to real revenue and productivity growth (i.e., real revenue per worker).

High growth firms are part of the ongoing dynamics of real revenue and input reallocation that characterize U.S. economic growth. Since at least the work of Dunne, Roberts and Samuelson (1989) and Davis and Haltiwanger (1990, 1992) we have known that underlying net growth in the U.S. is a high pace of job reallocation. Early work focused on decomposing net employment growth into gross job creation and destruction. More recent work has shown that there is a high pace of real revenue and capital reallocation that accompanies the employment reallocation (see, e.g., Foster, Haltiwanger and Krizan (2001), Becker et. al. (2005)) at least for selected sectors. One of the earliest findings in this literature is that young businesses exhibit a high pace of reallocation relative to more mature businesses. A second key finding in the early literature is that most of the job reallocation reflects reallocation within industry. While early work focused on U.S. manufacturing, recent work has extended the analysis to the entire U.S. private sector (e.g., Haltiwanger, Jarmin and Miranda (2013) and Decker, Haltiwanger, Jarmin and Miranda (2014)).[3]

The high pace of within industry reallocation has been interpreted in terms of the canonical firm dynamic models of Jovanovic (1982), Hopenhayn (1992) and Ericson and Pakes (1995) amongst others. In these models and in the subsequent literature, firms in the same industry differ in their productivity and the reallocation dynamics reflect moving resources away from less productive to more productive businesses. Such productivity differences can be endogenous given the role of endogenous innovation and R&D activities. Entrants and young

[3] Hereafter we often refer to these as HJM (2013) and DHJM (2014).

businesses play a critical role in these dynamics. They put competitive pressure on incumbents and in some models they are critical for innovation (see, e.g., Acemoglu et. al. (2013)).

The high pace of real revenue and input reallocation of young businesses is interpreted as part of the learning and selection dynamics as well as the endogenous innovation dynamics that are present in this class of models. Jovanovic (1982) argues that entering firms don't know their type and learn about it over time. In that model, high growth young firms are those that learn that they are high productivity or high demand. In contrast, high decline young firms are those that learn that they are low productivity or demand. Ericson and Pakes (1995) extended these learning ideas to environments where all firms engaging in some new form of activity have to learn whether they are profitable in that activity. Moreover, with endogenous innovation such as in Acemoglu et. al. (2013) productivity evolves based on the amount and success of innovative activity. In these models with more active learning and endogenous innovation, high growth young firms are those that innovate and learn successfully.

These models highlight the importance of understanding the contribution of and the nature of high growth as well as high decline businesses. Many issues are important in this context. It is now well understood that many distortions may arise in the ongoing growth and survival dynamics of firms (see, e.g., Hopenhayn and Rogerson (1992), Restuccia and Rogerson (2007), Hsieh and Klenow (2009, 2014), Bartelsman et. al. (2013)). In this respect, understanding the details of the life cycle dynamics of firms is critical since time periods, countries, geographic areas within countries, and industries with distortions may be identified by finding anomalies in the firm growth dynamics.[4] Moreover, it is increasingly understood that in some sectors the role of reallocation dynamics for growth may be much more important than

[4] For example, Hsieh and Klenow (2014) identify distortions in Mexico and India through the finding that young firms exhibit little net growth in Mexico and India compared to the US.

others. For example, the recent work of Hurst and Pugsley (2012) highlights the heterogeneity in the motivation for starting a business and hence their potential growth. They point to sectors dominated by small businesses that reflect occupational choices (such as wanting to be their own boss) rather than a desire to innovate and grow. In such sectors, it may be the case that high growth firms do not play a significant role in contributing to job creation and productivity growth.

In this paper, we make progress on these issues by characterizing the distribution of young firms by employment growth and also real revenue growth. Many previous efforts focusing on high growth firms have not studied how such firms fit into the entire distribution of firm growth rates. Moreover, having both real revenue and employment for firms permits measuring and exploring the connection between these dynamics and a measure of productivity growth.

The paper proceeds as follows. Section II presents a description of the data developed and used in this paper. Section III presents our main empirical findings. Our findings are mostly descriptive findings about the joint distribution of employment, real revenue and productivity growth. Given our interest in entrepreneurship, we focus considerable attention on the role of young firms in these dynamics. Concluding remarks that summarize our main findings and discuss next steps are in section IV.

II. Business Dynamics Data

We use two core related databases in this paper. Both are based on the Census Business Register (BR). We use the Census Bureau's Longitudinal Business Database (LBD) to construct measures of firm employment growth and firm age. Like the BR, the

LBD covers the universe of establishments and firms in the U.S. nonfarm business sector with at least one paid employee. The LBD includes annual observations beginning in 1976 and currently runs through 2012. It provides information on detailed industry, location and employment and parent firm affiliation for every establishment. Employment observations in the LBD are for the payroll period covering the 12th day of March in each calendar year. The LBD's high quality longitudinal establishment and firm linkages of establishments make possible the construction of our measures of firm growth and firm age. In what follows, we first discuss the key features of the LBD and then return to discussing the data we use from the BR to measure real revenue.

A unique advantage of the LBD is its comprehensive coverage of both firms and establishments. Firm activity is captured in the LBD up to the level of operational control instead of being based on an arbitrary taxpayer ID.[5] The ability to link establishment and firm information allows firm characteristics such as firm size and firm age to be tracked for each establishment. Firm size measures are constructed by aggregating the establishment information to the firm level using the appropriate firm identifiers. The construction of firm age follows the approach adopted for the BDS and based on our prior work (see, *e.g.*, Becker, et al. (2006), Davis, et al. (2007) and Haltiwanger, Jarmin and Miranda (2013)). Namely, when a new firm ID arises for whatever reason, we assign the firm an age based on the age of the oldest establishment that the firm owns in the first year in which the new firm ID is observed. The firm is then

[5] A closely related database at the BLS tracks quarterly job creation and destruction statistics (Business Employment Dynamics). The BED has advantages in terms of both frequency and timeliness of the data. However, the BED only can capture firm dynamics up to the level of establishments that operate under a common taxpayer ID (EIN). There are many large firms that have multiple EINs – it is not unusual for large firms operating in multiple states to have at least one EIN per state.

allowed to age naturally (by one year for each additional year it is observed in the data) regardless of any acquisitions and divestitures as long as the firm continues operations as a legal entity. This permits defining startups as new firms with all new establishments and shutdowns as firms that cease operations and all establishments shut down.

We utilize the LBD to construct annual establishment-level and firm-level employment growth rates. The measures we construct abstract from net growth at the firm level due to M&A activity. We use Davis, Haltiwanger and Schuh (1996) net growth rate measures that accommodate entry and exit.[6] We refer to this as the DHS growth rate.

Computing establishment-level growth rates is straightforward but computing firm-level growth rates is more complex given changes in ownership due to mergers, divestitures, or acquisitions. In these instances, net growth rates computed from firm-level data alone will reflect changes in firm employment due to adding and/or shedding continuing establishments. This occurs even if the added and/or shed establishments experience no employment changes themselves. To avoid firm growth rates capturing changes due to M&A and organization change, we compute the period *t-1* to period *t* net growth rate for a firm as the sum of the appropriately weighted DHS net growth rate of all establishments owned by the firm in period *t,* including acquisitions, plus the net growth attributed to establishments owned by the firm in period *t-1* that it has closed before period *t*. For any continuing establishment that changes ownership, this method attributes any net employment growth to the acquiring firm. Note, however, if the acquired establishment exhibits no change in employment, there will be no accompanying

[6] This growth rate measure has become standard in analysis of establishment and firm dynamics, because it shares some useful properties of log differences but also accommodates entry and exit (See Davis et al 1996, and Tornqvist, Vartia, and Vartia 1985).

change in firm-level employment induced by this ownership change. The general point is that this method for computing firm-level growth captures only "organic" growth at the establishment level and abstracts from changes in firm-level employment due to M&A activity (see supplementary data appendix to HJM (2013) for an example).

The LBD permits us to characterize the comprehensive distribution of firm growth rates including the contribution from firm entry, firm exit and continuing firms.[7] We begin our analysis below with the LBD to characterize the distribution of firm net employment growth rates for both continuing and exiting firms. Much of our analysis focuses on firms that are age 1 and greater so that we don't focus on startups in their first year. Our recent work (see, Haltiwanger, Jarmin and Miranda (2013)) highlights the contribution of startups to job creation in their first year. As we noted in the introduction, startups account for slightly more than 20 percent of firm level gross job creation (and slightly less than 20 percent of establishment level job creation). The focus of the current paper is post-entry dynamics.

As noted in the introduction, a key innovation of this paper is that we introduce real revenue and productivity growth measures to the analysis of high growth firms. Our measure of revenue is from the Census Bureau's Business Register (BR) which also provides the source data for the LBD. The BR's revenue measure is based on administrative data from annual business tax returns. Unlike payroll and employment, which are measured at the establishment level going back to 1976, the nominal revenue data are available at the tax reporting or employer identification number (EIN) level starting in the mid-1990s. Thus, in the BR, revenue is only measured for the establishment units of single location firms. Constructing a comprehensive revenue measure is further complicated by the fact that the content of the receipts fields on the

[7] By continuing firms we mean firms that continue between t-1 and t.

BR vary substantially by type of activity and the legal structure of the firm according to different tax treatments. Finally, the BR receipts fields are more likely to suffer from missing data than fields more critical to Census operations such as payroll and employment. These measurement issues are why revenue measures have yet to be incorporated into the LBD.

We successfully added nominal revenue measures to over 80 percent of the firm records in the LBD in our sample period. Given that we don't have complete coverage of the revenue data for all firm-year observation, we construct a sample of two-year pairwise continuing firms to allow computation of growth rates. That is, these are firms with both revenue and employment data in t-1 and t. This group is a subset of the continuing firms in the full LBD. We describe this in more detail in the data appendix.[8] Although given the structure of the data selection issues are possible, we have found that the pattern of missingness of revenue for the two year pairwise continuers is only weakly related to the measures of firm age, firm size or industry that we have from the LBD as described above. As such, we think it is reasonable to treat revenue as approximately missing at random in terms of the co-variates of interest here.[9] The justification for this treatment of the data will become apparent in the next section since we show that the restricted database yields patterns quite similar to the universe for employment growth rates.[10]

We deflate the nominal revenue measures with a general price deflator (the GDP Implicit Price Deflator). As such, our measures of real revenue will reflect both real revenue changes and changes in relative prices across industries. Revenue fields in the BR can be noisy so we adopt

[8] The data appendix will be added in the next draft of the paper.

[9] In unreported results that we plan to incorporate in future drafts, we have used inverse probability weighting (IPW) (see Wooldridge (2002, 2006)) to address the modest systematic variation in missingness we have found by industry, firm size, firm age and multi-unit status. The use of such weights can reduce the small discrepancies we find in Figures 3 and 4 between the full LBD and our revenue restricted subset.

[10] We note that we exclude 2001 and 2002 from our statistics since the 2001 data are problematic (which impacts the growth rate distributions in both 2001 and 2002).

filters to clean out unreasonable values. These filters are discussed further in the data appendix and include minimum and maximum productivity value cutoffs, maximum revenue cutoffs, and maximum revenue growth values. Subsequent references to revenue in what follows should be interpreted as real revenue. When we examine labor productivity, we focus on within industry patterns since these measures are not comparable across industries. We also note that for revenue growth we use DHS measures of growth. One potentially important limitation of our revenue growth measures is that since we don't have the underlying establishment level revenue growth we cannot abstract from the contribution of M&A activity to revenue growth. The filters we design partly take care of this as M&A activity can lead to spurious large flows of revenue. We have checked and found that the broad patterns we find for employment growth largely hold when we do not adjust for M&A growth – but still we regard this as a limitation that should be acknowledged (and also and also as an area for future research).

III. The Role of High Growth Firms for Job Creation, Real Revenue Growth and Productivity Growth

A. The Up or Out Dynamics of Young Firms in the U.S.

We begin by comparing results we obtain with the revenue enhanced subset of the LBD with prior findings from HJM and DHJM which make use of the full LBD. Those papers emphasized two features of the employment growth dynamics of young firms in the U.S. The first is the up or out dynamic of young firms. Figure 1 shows the net growth rate for surviving firms as well as the job destruction rate from firm exit by firm age for the full LBD except we exclude years not covered by the revenue enhanced subset[11]. Firm exit is defined as discussed above. All statistics are employment-weighted. Figure 1 focuses on the post entry dynamics of firms; in our nomenclature, age one is the year after entry. We exclude entrants in this figure since age zero businesses inherently only create jobs in their year of entry.[12] The weighted sum of net job creation yields overall net employment growth for a given age group.[13] Conditional on survival, young firms have much higher growth rates than more mature firms. Young firms also have a substantially higher (employment weighted) exit rate than more mature firms. Fifty percent of an entering cohort of firms will have exited by age five (on an employment weighted basis). The very high failure rate of young firms is partially offset by the contribution of the surviving firms. For the sample period in Figure 1, five years after the entry of an average cohort, the employment is about 70 percent of the original contribution of the cohort. This is in spite of losing about fifty percent of employment to business exits.[14]

One implication of Figure 1 is that the overall net employment growth rate is negative for all firm age groups. This pattern is evident from the job destruction rate from exit exceeding the net employment growth rate for continuing firms for all firm age groups. This pattern partly

[11] In particular, the statistics are based on tabulations of pooled data from 1998-2011 from the Longitudinal Business Database (LBD) excluding the 2001 and 2002 years. We exclude those years since here since the revenue data for those years are not of sufficient quality to generate robust growth rate distributions. As we discuss below, the focus on the 1998-2011 period implies that our statistics are influenced substantially by the Great Recession.

[12] See HJM (2013) and DHJM (2014) for an extensive analysis of the contribution of startups to job creation. We have noted their average contribution. Those papers highlight that there has been a declining pace of entry in the U.S. They also note that entry rates vary substantially across sectors and geographic regions. But interestingly the papers note even with variation in the entry rates that the post-entry dynamics are similar across sectors in terms of up or out dynamics. It is the latter that we focus on in this paper and indeed focus mostly on the “up” component. Of course, one cannot forget the out component as well. We view the latter as part of the volatility that we discuss at length throughout the paper.

[13] Overall net growth is the sum of the weighted net growth rate for continuers plus job destruction from exit. The weight is the share of employment for continuing firms. See HJM (2013) for details.

[14] These calculations of the five year contribution of each cohort are low relative to those reported in HJM (2013) or in DHJM (2014). These differences reflect differences in sample periods and in particular whether the years of the Great Recession are included. HJM (2013) use the period 1992-2005. They find that for five years after the entry of an average cohort, the employment is about 84 percent of the original cohort. DHJM (2014) use the period 1992-2011 and find the same calculation yields 80 percent.

reflects our sample period which includes the sharp contraction and slow recovery of 2007-11. But it also reflects the more general pattern that even in a typical year of overall positive net growth, continuing firms tend to be mildly contracting on average with overall (economy wide) net employment growth being positive because of the contribution of firm startups (not depicted in Figure 1). HJM show that this pattern holds for the sample period 1992-2005.[15] A related implication of Figure 1 is that overall net employment growth rates are increasing with firm age.[16] Again, this partly reflects our sample period since young firms were hit especially hard in the Great Recession (see, Fort et. al. (2013)) but is also a common pattern more generally (see Figure 4 of HJM). As highlighted in HJM and as illustrated in Figure 1, we think it is instructive to decompose overall net growth into the net growth from continuers and the contribution from exit. It is via this decomposition that the up or out pattern of young firms emerges. In turn, it is instructive to examine the underlying full distribution of firm growth rates underlying the patterns for continuing firms to which we now turn.

The second finding by DHJM highlights the dispersion and skewness of the employment growth rate distribution of continuing young firms. Figure 2 shows the 90th, 50th (median) and 10th percentiles of the net job growth distribution of surviving firms by firm age. Percentiles are from the employment-weighted distribution which mitigates the impact very small firms have on these statistics. We discuss dispersion by examining the patterns of the 90-10 differential and skewness by comparing the difference between the 90-50 and the 50-10 differentials.

Young continuing firms have very high dispersion of employment growth, and also very high positive skewness. The median employment growth rate for young firms is close to zero

[15] The BDS shows that in the years of most robust net growth, both very young and very old firms tend to have positive overall net growth inclusive of the contribution of exit.

[16] This can be inferred by computing the overall net growth implied by Figure 1.

(and for that matter the median is close to zero for all firms) so the positive skewness is seen in the relative magnitudes of the 90th and 10th percentiles where the employment growth rates of younger firms are much more skewed to the right (positive) compared to more mature firms. This accounts for the high mean net employment growth rate of young firms relative to older firms we see in Figure 1. Taking Figures 1 and 2 together, the typical young continuing firm (as captured by the median) exhibits little or no employment growth even conditional on survival; however, amongst all the young firms a small fraction exhibit very high rates of growth.

Keeping the pattern in figures 1 and 2 in mind, we characterize the distribution of revenue growth rates. Recall that we were unable to measure revenue for all firm-year observations in the LBD. Thus, our revenue enhanced subset of the LBD consists of two-year pairwise continuing firms. This subset comprises more than 80 percent of the firms in the LBD for the years 1998-2000 and 2003-2012.

Figure 3 shows the relationship between the net employment growth rates for continuing firms by firm age for the full LBD and for our revenue enhanced subset. Our restricted subset yields very similar patterns to the full LBD for continuing firms.[17] Figure 4 shows the 90th, 50th and 10th percentiles of the employment-weighted distribution of firm net employment growth rates for our restricted sample of continuing firms. Comparing Figure 4 and Figure 2 shows that the full LBD and our restricted sample again yields very similar patterns for continuing firms. Figures 3 and 4 give us the confidence to proceed with our revenue enhanced subset of continuing LBD firms for the remainder of the analysis. As we discussed in section II, our focus on the remainder of the analysis is on surviving firms. We know from Figure 1 as well as the

[17] The magnitudes in Figure 3 are sensitive to the inclusion of the Great Recession years. For example, the mean net growth rate for continuing one year old firms in 1998-06 is about 17 percent but only about 10 percent for 2007-11. The differences are larger for young firms. For example, for 16+ year old firms the mean net growth rate is about 1 percent for 1998-06 and -1 percent in 2007-11.

extensive analysis of firm exit in the literature that young and small firms have very high rates of exit. Such exit patterns need to be kept in mind to provide perspective on the results we focus on.

B. Real Revenue vs. Net Employment Growth Rate Distributions by Firm Age

Figure 5 characterizes the distribution of firm revenue growth rates by firm age for continuing firms. Depicted are the 90th, 50th, and 10th percentiles of the revenue weighted distribution. As before, weighting mitigates the impact that very small firms have on these statistics since they account for only a small fraction of revenue. Comparing Figures 4 and 5 yields some similarities, but also some notable differences. First, consider the median revenue growth vs. median net employment growth rate patterns by firm age. A notable difference is that the net employment growth rate for the median surviving firm is close to zero except for age one firms where it is about 5 percent. In contrast, the median surviving firm has positive revenue growth in excess of 4 percent per year in each of the first four years and in excess of 2 percent in all years. Second, the 90-10 differential is substantially larger for young firms compared to more mature firms. This is especially true for net employment growth but also holds for revenue growth in periods of economic expansion.

Third, as we have noted above, young firms exhibit considerable skewness in net employment growth rates. This skewness is, perhaps surprisingly, less apparent for revenue growth than it is in the case of employment growth. Although, cyclical dynamics aren't our primary focus, it's important to note that cyclical factors matter for the skewness of these growth rate distributions, but especially so for revenue growth. Figures 6a and 6b depict the 90-50 and 50-10 differentials for revenue growth (6a) and net employment growth (6b) for the sub-periods 1998-06, prior to the recession, and 2007-11, recession and post-recession. The cycle clearly influences the skewness patterns. In Figure 6a, we find that the 90-50 exceeds the 50-10 for revenue growth for all firm ages at or below 5 and that the 90-50 is about the same as the 50-10 for firm ages greater than 5 for the period 1998-06.[18] However, in the latter (recession and recovery) period, the 50-10 differential increases substantially for all ages so that rather than positive skewness, the revenue growth distribution exhibits negative skewness for most ages and especially for older firms. Figure 6b shows similar but more muted patterns. In short, we find that the positive skewness for young firms exhibited in terms of both revenue and net employment growth is procyclical. In what follows, for the sake of brevity, we will mostly present results for our entire sample period but we will note when patterns are especially sensitive to the business cycle.

Turning to Figure 6c, we find that the mean revenue and net employment growth rates for surviving firms decline sharply with firm age. Based on Figures 4, 5, 6a and 6b, we know that underlying these quite similar mean patterns are differences in the shapes of the underlying distributions. For net employment growth, the high mean for young firms is driven by the positive skewness for young firms. Or put more simply, the high average is driven by high growth firms. For revenue growth, the high mean for young firms reflects both the high median for young firms and the greater positive skewness for young firms.

In either case, Figures 4 and 5 highlight the very high net employment and revenue growth of the 90th percentile firms particularly for young firms. We quantify their importance in Table 1 where we decompose revenue and employment growth. We find that13 percent of continuing firms have revenue growth in excess of 25 percent accounting for about 50 percent of

[18] The exclusion of 2001 and 2002 from our 1998-06 sample period may be playing a role here as well. However, we note that the full LBD shows substantial positive skewness in the employment growth rate distribution using all years from 1998-06.

the gross revenue creation for continuing firms.[19] Analogously, about 17 percent of continuing firms have net employment growth in excess of 25 percent accounting for close to 60 percent of gross job creation for continuing firms.

In what follows, we explore the characteristics of high growth firms on a number of margins. In particular, we consider not only firm age, but firm size, industry and geographic location. We turn to that analysis below. Before doing so, we provide evidence on the connection between revenue and net employment growth rates.

C. The Joint Distribution of Real Revenue and Net Employment Growth Rates.

Theoretical models of firm adjustment in response to shocks suggest a positive correlation between revenue and employment growth. This correlation may depend on the nature of the adjustment costs and frictions. We find that revenue and net employment growth rates for surviving firms are positively correlated but the correlation is not high (about 0.22). Further analysis shows that this reflects in part the pattern that revenue growth rates tend to lead employment growth rates. Table 2a shows the estimates from a simple reduced form one lag VAR model relating firm level net employment growth and revenue growth for continuing firms.[20] Net employment growth estimates reported in the first column show there is negative serial correlation reflecting the well-known regression to the mean in employment growth rates. Interestingly, however, lagged revenue growth is associated with higher net employment growth in the current period. The same is not true (to the same extent) for the relationship between lagged net employment growth and current period revenue growth, shown in column 2,

[19] By this we mean, the revenue creation from growing firms.
[20] We weight the regressions with the LHS employment growth with employment weights and the regressions with the RHS revenue growth revenue weights. We have tried common weights and obtain similar results.

suggesting first, that the revenue shock leads the employment adjustment, and second, that revenue growth is only weakly correlated with prior growth shocks.

The patterns in Table 2a are consistent with standard adjustment cost models for employment dynamics (see, e.g., Cooper, Haltiwanger and Willis (2007)).[21] In such models, firms facing a positive profit (e.g., demand or productivity) shock exhibit immediate increases in revenue but a delayed adjustment for factors such as capital and labor.

We now explore whether the patterns at the mean of the growth rate distributions carry over to the upper tails of the joint growth rate distribution.   Table 2b shows results for a similarly estimated simple one lag VAR models for indicators for high growth episodes firms. For this purpose, a firm experiences a high revenue (employment) growth episode in a particular year if the firm's revenue (net employment) growth rate is greater than 25 percent.

Table 2b shows that having a high growth revenue episode in the previous year is positively associated with having both high revenue and employment growth episodes in the current year. Interestingly, in spite of the overall negative serial correlation for employment growth in Table 2a, there is some positive persistence in high employment growth episodes. These patterns are consistent with high revenue (employment) growth events extending beyond a single year, with high revenue growth events tending to precede high employment growth events.

D. Within Firm High Growth and High Decline Events: Persistence vs Volatility.

We now focus our attention on the characteristics of high growth firms defined, as above, as firms with annual growth in excess of 25 percent.   In exploring the characteristics of high

[21] This likely also reflects the timing of the data. Employment growth from t-1 to t represents a March-to-March change while revenue growth represents annual revenue changes from t-1 to t. Our primary focus is not on dynamics so we don't explore this issue further.

growth firms, we recognize that the propensity for a firm type to be high growth may be capturing variation related to first moment, second moment, and third moment effects of the growth rate distribution. That is, the distribution of firm types with a high propensity to produce high growth firms are likely to have high means, high dispersion and positive skewness. As already seen in Figures 3 and 4, the growth distributions of young firms have high means, high variances and high positive skewness relative to more mature firms. The high variance implies that there will be some tendency for groups, such as young firms, with a high fraction of high growth firms to also have a high fraction of high decline firms. Put differently, it may be that the characteristics that we identify as being associated with high growth are, in part, identifying firm types that are highly volatile. But our results on firm age imply that a greater propensity for high growth is associated with more than just volatility. Young firms are more likely to be high growth and also have higher mean growth driven in part by greater skewness. As we explore these characteristics, we will explore the connection to first moment, second moment and third moment effects.

Before proceeding, it is helpful to examine the propensity for high growth events in a given firm to be associated with high decline events within the same firm. That is, how often are high growth events reversed? In a related fashion, it is useful to examine the propensity for repeated high growth and high decline events within the same firm. Figure 7a illustrates the joint distribution of high growth and high decline (contracting by more than 25 percent) episodes within firms that are five years old on a revenue growth basis. For each five year old firm, we count the number of high growth and high decline events that the firm has experienced. A five year old firm can have between 0 and 5 high growth and high decline events. Of course, the total number of high growth and high decline events can only add up to 5. The joint distribution is characterized using the average revenue of each firm as weights.

For reference, 61 percent of revenue is at five year old firms that have zero high decline events and 40 percent of revenue is at firms that have zero high growth events. 27 percent of revenue is at five year old firms having no high growth or high decline events. Unconditionally 28 percent of revenue is at firms that have one high growth event, 19 is at firms that have two high growth events, 10 percent is at firms that have three high growth events, 4 percent is at firms that have four high growth events, and 1 percent is at firms that have 5 events. The respective figures for high decline events are 26, 10, 3, 0 and 0 percent. Comparing the two indicates the asymmetry that underlies the positive skewness. Conditional probabilities (revenue weighted) are easily computed from this joint distribution. The probability that a five year old firm with one, two, three and four high growth events has zero high decline events is 54 percent, 50 percent, 59 percent and 74 percent respectively.[22] Thus, most five year old firms with one or more high growth events have no high decline events. In that respect, high growth firms are not simply an indicator of high within firm volatility. Still there are firms with both high growth and high decline events. For such firms, high growth events are also associated with high within firm volatility.

Figure 7b provides the same type of joint distribution on an employment growth basis with the shares depicted reflecting the share of employment in each of the cells. The overall shape of the joint distribution is similar to that for revenue growth. Again, we find that the five year old firms with one or more high employment growth events are more likely to have experienced no high decline events on an employment weighted bases. For example, 61 percent

[22] The (revenue weighted) probability that a firm with zero high growth events has zero high decline events is about 70 percent. This is not surprising since most revenue is at firms with zero high decline events.

of employment is at five year old firms with one or more high growth employment events. 47 percent of the employment at those firms is at firms with zero high employment decline events.

E. *The Characteristics of High Growth Firms: By Firm Age, Firm Size, Industry and Geographic Location*.

Our objective is to provide descriptive statistics about the characteristics of the types of firms that find themselves in the top of the growth rate distribution. To this end, we estimate linear probability regressions pooling across firm years. We consider discrete dependent variables that take on a value of one if the firm is a high growth revenue (employment) firm. As before, we define high growth firms as those with annual growth in excess of 25 percent.[23] For the specifications with high revenue growth indicators we weight by revenue (averaged in period t-1 and t) and for the specifications with high employment growth indicators we weight by employment (averaged in t-1 and t).

We first focus on firm age and firm size characteristics. For firm age, we consider firm age classes between 1 and 16+. For firm size, we use within industry deciles of the size distribution. In the case of the revenue growth specifications, these are revenue weighted deciles of revenue size. For the employment growth specifications, we use employment weighted deciles of employment size. For calculating these deciles, we use two alternative measures of size for revenue and employment. We use base year size (e.g., revenue or employment in period t-1) and current average size (i.e., the average of revenue or employment in period t-1 and t). We consider both since as discussed in HJM using base year size yields regression to the mean effects (i.e., given transitory shocks a firm classified as small in the prior period is more likely to grow). The use of current average size is a compromise between using base year and current

[23] There is nothing inherently special about the 25 percent cutoff. We have found our results are robust to using alternative cutoffs.

year size (where the latter suffers from the opposite problem from base year size). We present our estimated firm size and firm age coefficients via a series of graphs. We don't report standard errors but note given the very large sample size (in excess of 30 million) all of the standard errors for the reported size and age effects are less than 0.001. The same remarks apply to the state and industry effects that we report below.

Figures 8 and 9 report the estimated firm age effects for high growth employment and revenue firms respectively with and without size controls. The likelihood of being a high growth employment and revenue firm is decreasing with firm age even with firm size controls. The latter have relatively little influence on the patterns. It is apparent that our earlier findings in Figures 2 and 4 are robust to controlling for firm size effects. We note that in unreported results we also find that these patterns are robust to controlling further for industry and year effects.

Figures 10 and 11 report the analogous estimated firm size effects for high growth employment and revenue firms with and without age controls. For the firm size effects, we report results using both base year and current average size categories. If we don't control for firm age, there is an inverse relationship between firm size and the likelihood of being a high growth firm using both the base year and current average size approaches. But once we control for firm age, these patterns are substantially mitigated. For high employment growth firms, the relationship between the likelihood of being a high growth firm and size is relatively flat using current year average size and age controls. For high revenue growth firms, the relationship becomes partly positive.

The inference we draw from Figures 8-11 is that firm age is a robust and key determinant of the likelihood of being a high growth firm. In contrast, once we control for firm age, firm size has relatively little influence. The role of firm age as opposed to firm size is reminiscent of the findings in HJM that found that young firms grow faster than more mature firms but that small firms do not grow faster than large firms once firm age is taken into account. We note, however, that while firm age is a key determinant that the adjusted R-squared from age effects alone is 2 percent for the revenue growth distribution and 5 percent for the employment growth rate distribution. With size, industry, state and year effects the adjusted R-squared rises to about 9 percent for revenue growth (using either base year or current average year size) and between 8 and 9 percent for employment growth. Industry effects alone yield 7 percent and 4 percent respectively in terms of adjusted R-squared. These patterns imply that the factors that determine which firm is a high growth firm largely are factors within firm age, firm size, industry, and year cells that we do not observe in our data. Still there is systematic variation by industry and state to which we turn to now.

Figure 12a shows the top fifty industries and Figure 12b shows the bottom fifty industries for high revenue growth firms. Figures 13a and 13b show the analogous patterns for high employment growth firms. Reported are the regression estimates with industry effects alone. We begin by noting that all 4-digit NAICS sectors have some high growth firms. The top ranked industries have high growth firms that account for as much as 45 percent of industry revenue and 30 percent of industry employment. In contrast, the bottom ranked industries have high growth firms that account for less than 1 percent of industry revenue and employment.

The patterns for high growth by revenue and employment are positively correlated as can be seen in Figure 14 but the correlation is far from one. This is apparent from an examination of the industry lists. The top industry for high growth firms in terms of revenue is the oil and gas extraction industry (NAICS 2111). This industry does not make the top fifty for high growth firms in terms of employment. The top industry for high growth firms in terms of employment is telecommunications resellers (NAICS 5173) which is in the top 50 (at rank 13) for revenue. Other high tech sectors make the top 50 of both lists. For example, internet service providers and portals (NAICS 5181) is second on the high growth firm list for employment and 8th for revenue.

Looking at the bottom 50 industries, there is again some overlap but also some distinct differences. Motor vehicle manufacturing (NAICS 3361) and tobacco manufacturing (NAICS 3122) are towards the bottom in terms of both high revenue growth and high employment growth firms. But iron and steel mills (NAICS 3311) is towards the top of the revenue rankings and towards the bottom of the employment rankings. It is also not apparent to us that the bottom industries for either revenue or employment are inherently dominated by sectors where being one's own boss is especially attractive. In this respect, the hypothesis of Hurst and Pugsley (2012) that some industries are dominated by such businesses is not apparent in these distributions.

Table 4 reports analysis of whether there are industry clusters that are more or less likely to have high growth firm activity. The industry clusters we consider are sectors that can be classified as tradable, construction, high tech, bio tech and energy related. We find that the energy related sectors have greater high growth firm activity in terms of revenue. For high tech we find positive estimates that are not statistically significant. Tradable are much less likely to have high firm growth activity in terms of employment. The latter is consistent with the view that employment gains from tradable have largely been off-shored in our sample period.

The stability of the industry rankings over time is depicted in Figure 15. This figure shows the industry rankings in the 1998-06 period relative to the industry rankings in the 2007-11 period. Industries are clustered along the 45 degree line but there are some notable exceptions. Industries with fewer high growth firms are located nearer the origin. Thus, industries moving up the rankings are located in the upper left portion and those moving down are in the lower right. While we have not rigorously analyzed these changes, some interesting findings arise even from a quick review of figure 15. First, industries such as food processing that have seen high demand and prices fueled by exports have seen increases in the number of high growth firms. Conversely, industries affected by recession such as financial and real estate services saw decreases in the number of high growth firms.

We have already determined that high growth events can be reversed especially for highly volatile firms. We have also determined that high growth events can persist accounting for the lasting impact a new cohort of young firms has on growth. Table 5 presents pairwise correlations between the high growth firm industry estimates and summary statistics of the first, second and third moments of the within industry distributions. We find that industries exhibiting greater high growth firm activity also tend to have high overall mean growth on both a revenue and employment basis. But industries with a large fraction of high growth activity also have a large fraction of high decline activity and greater volatility as captured by the 90-10 differential. In the case of industries exhibiting greater employment based high growth activity, they also exhibit greater positive skewness in the employment growth rate distribution (as measured by the difference between the 90-50 and 50-10 differentials in net employment growth rates).

To further explore the relationship between overall growth and high growth firm activity, Figure 16a shows the scatter plot relating industry high growth firm effects against industry overall revenue growth rates. Figure 16b shows the analogous pattern for employment growth. The strong positive correlation between overall growth and high growth firm activity is evident for both revenue and employment. There are some interesting outliers for revenue growth that take the form of high overall growth without substantial high growth firm activity. For example, pipeline transportation of crude oil (NAICS 4861), interurban and bus transportation (NAICS 4852), and taxi and limousine service (NAICS 4853) all are industries with reasonably high overall growth (especially NAICS 4861) but not much high growth firm activity. These are industries that arguably have had robust growth without the type of volatile high growth young firm activity that appears to be important for other sectors.

We repeat the same type of exercise for state effects in Figures 17 and 18 for high growth real revenue firms and high growth employment firms respectively. States at the top on the basis of revenue growth are energy intensive states such as Oklahoma, Texas and North Dakota. For the ranking by high growth firms by employment, Oklahoma is towards the top but Texas and North Dakota are not. In spite of the above exceptions, Figure 19 shows that there is greater cross sectional correlation in high growth revenue and employment states compared to industries. Figure 20 examines the stability of the rankings and shows somewhat less stability of state rankings in terms of more departures from the 45 degree line. States such as North Dakota have increased their ranking in terms of high growth revenue firms while states like Arizona have decreased their ranking.

Table 6 shows the correlations for the state indicators of the mean, dispersion (90-10) and skewness (90-50 minus 50-10). Here again we find that high growth firm effects for a state are positively related to overall growth, dispersion and skewness. Figures 21a and 21b show the evident strong relationship between overall high growth and high growth firm activity by state. Unlike the industry analysis, there are no obvious outliers.

IV. Firm Age and Productivity Dynamics: The Role of High Growth Young Firms?

We now turn to the relationship between high growth young firms and productivity dynamics. This is the first economy wide database including measures of revenue and productivity on an annual basis. We are interested in evaluating the role of reallocation in productivity dynamics first and then specifically on the contribution that high growth young firms have on the reallocation components of productivity growth. We use the revenue and employment measures to construct a labor productivity measure for each firm. Since we use gross revenue and not value added, these statistics are not comparable across industries so we again focus on within industry patterns. In addition, we don't readily have available industry specific revenue deflators for all industries such as those available for manufacturing industries.

We begin with some cross sectional analysis where we abstract from between industry variation as well as variation over time. Specifically, for this cross sectional analysis we deviate firm level (log) revenue per worker from its industry-level mean in each year.[24] We then explore the variation in first and second moments in these relative productivity measures across firm ages. Figure 22 shows the mean and standard deviations of the within industry (log) labor productivity measure by firm age. We construct this figure as follows. First, we compute the within industry means and standard deviations within each detailed 6-digit industry for each firm age group. In the top panel (22a), we generate these means on an unweighted basis. In the bottom panel (22b), we use employment weights to weight up to the industry level. Then in both the top and bottom panels we take an average across all industries where we use gross revenue weights following the procedures used in Foster, Haltiwanger and Krizan (2001) and Baily, Hulten and Campbell (1992). For the mean calculation we index the average productivity of 16+ year olds at 1 so that the reported effects reflect differences from that oldest group.

[24] As such, this exercise effectively controls for industry-specific prices.

Figure 22a shows that, relative to other firms in the same industry, mean (log) labor productivity rises with firm age whether we use the unweighted or weighted approach within industries. However, the differences by firm age are much larger in magnitude using the weighted approach. When we weight by employment, the patterns reflect both the unweighted mean within the industry, firm age cell and the covariance between size and productivity within the cell as per the Olley and Pakes (1996) decomposition.[25] The weighted mean patterns show a more dramatic increase with firm age. By construction, this pattern reflects a sharp rise in the covariance between size and productivity within an industry by firm age cell. The latter pattern is not surprising since for young firms the relationship between size and productivity is likely weak as firms have not sorted themselves out in terms of the relationship between relative size and productivity. Another possible factor is measurement error is greater for young firms but this should be less problematic in this setting given the use of administrative data.

Figure 22 also shows that the within industry dispersion of productivity declines monotonically with firm age. For both the unweighted and weighted results, we find similar patterns. The patterns are consistent with our findings of much greater dispersion in both revenue and net employment growth for young firms.

To explore the contribution of high growth firms, we turn to examining within industry decompositions of industry level productivity growth on continuing firms.[26] Here we start with

[25] The Olley-Pakes (1996) decomposition of the level of productivity is given by:

$P_{it} = \sum_{e\epsilon i} \omega_{et} P_{et} = \widetilde{P_{et}} + \sum_{e\epsilon i}(\omega_{et} - \widetilde{\omega_{et}})(P_{et} - \widetilde{P_{et}})$

Where a tilde represents the simple average across all plants in the same industry. When we compute the weighted average productivity for each age group and compare it to the unweighted average the difference is the Olley-Pakes covariance term for the age group.

[26] From the existing literature (see, e.g., Foster, Haltiwanger and Krizan (2001, 2006)), we know that net entry contributes disproportionately to within industry productivity growth as exiting businesses are much lower productivity than entering businesses. We are focusing on continuing firms given the limitations of our revenue restricted database.

firm level revenue per worker without deviating from industry or year effects. We control for industry and time effects via the decomposition itself. We define an index of industry level productivity as given by:

$$P_{it} = \sum_{e \epsilon i} \omega_{et} P_{et}$$

Where *e* indexes firms, *i* indexes industry, *P* is log labor productivity, ω is the share of employment. Note that for the purposes of a labor productivity index the appropriate weight is employment since then the index is the geometric mean of firm-level labor productivity. Then the change in this index at the industry level (which is log based so that it can be interpreted as an index of industry level productivity growth) can be decomposed into within and between effects as given by:

$$\Delta P_{it} = \sum_{e \epsilon i} \overline{\omega}_e \Delta P_{et} + \sum_{i} \overline{(P_e} - \bar{P}_i) \Delta \omega_{et}$$

where a bar over a variable represents the average over t-1 to t. The first term on the RHS is the within term and the second term is the between term. The within term captures the weighted average of within firm productivity growth while the between term captures the contribution of changes in employment shares. A firm contributes positively to the between term if it has labor productivity higher than the industry average. In this decomposition, we focus on within industry patterns by using an industry specific decomposition.

We calculate this decomposition for every industry, year pair in our data. To compute an aggregate average we use average gross revenue weights following the approach of Foster, Haltiwanger and Krizan (2001) and Baily, Hulten and Campbell (1992). Figure 23 shows the results of this decomposition for the average (revenue weighted) industry for all years and for the sub-periods 1998-06 and 2007-11. We find that the within term is highly procyclical. It is positive for the overall period, much higher in the 1998-06 period and negative in the 2007-11 period. This is consistent with within firm productivity being procyclical – likely for reasons associated with varying capacity utilization and the adjustment costs discussed earlier. In contrast, the between term is more stable over time and is always positive. It is not surprising then that overall and for the particular time period we are examining that the between term accounts for most of the overall increase for the full period and about half for the period 1998-06.

To explore the role of high growth firms, we focus on the between term since it is both more stable but also captures the reallocation dynamics where high growth firms play such a critical role. Figure 24 shows the contribution of each of the revenue growth rate classes (upper panel) and employment growth rate classes (lower panel). Interestingly, we find that it is especially high growth and high decline businesses that account for the between term and the patterns are roughly similar for both revenue and employment growth rate classes. For the growth rates classes with relatively modest increases or decreases we find little contribution of the between. Since the between term is only positive for a group of growing (shrinking) firms if they have productivity that is on average higher (lower) than the overall average, these findings imply rapidly growing firms have above average productivity while rapidly shrinking firms have below average productivity. In this respect, Figure 24 reminds us that the high growth firms are part of the overall dynamic contributing to productivity enhancing reallocation with an equally important role for rapidly shrinking low productivity firms.

Where do young high growth firms fit into this picture? First, we note that young firms that are less than 10 years old only account for about 13 percent of revenue and 19 percent of employment. But we find that that young firms contribute about 50 percent to the between term

– much higher than their share of activity. Also, high growth young firms contribute about 40 percent to the high growth component of the between term. Thus, we find that young firms disproportionately contribute to the between term overall and that high growth young firms contribute disproportionately to the between term contribution of high growth firms.

V. Concluding Remarks

We find that high growth young firms contribute disproportionately to job creation, revenue growth and labor productivity growth. Young firms are very heterogeneous. Many fail in their first few years and even amongst those that survive there is considerable dispersion. But conditional on survival, young firms have higher average net employment growth and revenue growth than their more mature counterparts. For employment growth, this is especially striking since median net employment growth for young firms is about zero. As such, the higher mean reflects the substantial positive skewness with a small fraction of very fast growing firms driving the higher mean net employment growth. For revenue growth, young firms have higher median growth than their more mature counterparts. Still, young firms exhibit more positive skewness in growth rates than their mature counterparts on both an employment and revenue growth basis – although (not the focus of the current paper) the positive skewness of revenue growth for young firms is highly procyclical.

Given these findings, we explored the characteristics of the high growth firms further. Consistent with the above, we find that high growth firms are more likely to be young than mature even controlling for firm size. We also found that there is considerable variation across industries and states in the fraction of activity accounted for by young firms. The range across industries and states is substantial. Industries at the top of the ranking have as much as 40 percent of activity in high growth firms while industries at the bottom of the ranking have close to zero. Differences across states are less substantial but still large. States at the top of the ranking have close to 20 percent of activity in high growth firms while states at the bottom have about 5 percent.

We find that the ongoing reallocation dynamics of which high growth young firms play a critical part contributes substantially to within industry labor productivity growth. Our findings suggest that at least half of within industry labor productivity growth is attributable to employment being reallocated from less productive to more productive firms within the industry. Young firms contribute disproportionately to this contribution from reallocation. But in this respect both high growth and high decline firms contribute substantially to the productivity enhancing reallocation.

For firm types with a greater fraction of high growth firms (e.g., young firms, specific industries, specific states), these types exhibit high overall net growth, higher volatility and also higher positive skewness. In this respect, a propensity for high growth is an indicator related to first, second and third moments of the growth rate distribution.

We interpret the statistical patterns as being consistent with models of innovation and growth that impact the first, second and third moments. A rough storyline that we think fits the patterns we have detected is as follows. Positive shocks and or endogenous innovative activity ultimately leads to growth but through complex dynamics involving both dispersion and positive skewness. The latter reflects the rareness of being a successful innovator and those that do succeed exhibit rapid growth. Those rare rapidly growing firms contribute substantially to net job creation, revenue growth and labor productivity growth. But often accompanying growth are those that do not succeed so that volatility accompanies growth.

This storyline is obviously incomplete on many dimensions. It may be that (and we have presented some evidence of this) that shocks and innovation in some sectors don't involve this complex dynamics of entry, exit, volatility and skewness. Another set of issues relate to industry life-cycle; that is, what do the dynamics of industries and locations in decline look like. They may also involve volatility. Similarly, there may be shocks that induce reallocation without much productivity growth or even adverse consequences for growth. For example, uncertainty shocks of the type emphasized by Bloom (2009) may have this character.

Our analysis has been intentionally descriptive. We think the data infrastructure we have developed and the basic facts we have presented provide a framework for more direct analysis of the process of innovation and growth. Our findings suggest that exploring patterns by firm age and examining first, second and third moment effects will be important for detecting and understanding periods of growth and innovation. Moreover, we think our data infrastructure and approach should be helpful to explore factors that distort innovation and growth. The recent findings of Hsieh and Klenow (2014) that show that young firms grow rapidly in the U.S. relative to their counterparts in India and Mexico is highly relevant in this context. Our findings show that the rapid growth of young firms in the U.S. involves substantial skewness and dispersion. As such, distortions that may be adversely impacting the growth of young firms in India and Mexico (amongst other countries) may be impacting many of the different margins that underlie the patterns we have detected.

References

Acemoglu, Daron, Ufuk Akcigit, Nicholas Bloom, and William R. Kerr. 2013. "Innovation, Reallocation, and Growth." NBER Working Paper No. 18993.

Bartelsman, Eric, John Haltiwanger and Stefano Scarpetta. 2013. "Cross Country Differences in Productivity: The Role of Allocative Efficiency," *American Economic Review,* 103(1): 305–334.

Becker, Randy A., John Haltiwanger, Ron S. Jarmin, Shawn D. Klimek, and Daniel J. Wilson. 2006. "Micro and Macro Data Integration: The Case of Capital." Chap. 12 in *Studies in Income and Wealth* edited by Dale W. Jorgenson, J. Steven Landefeld, and William D. Nordhaus, vol. 66, *A New Architecture for the U.S. National Accounts*, pp. 541-609. Chicago, IL: National Bureau of Economic Research, University of Chicago Press.

Bloom, Nicholas Bloom. 2009. "The Impact of Uncertainty Shocks," *Econometrica*, 77(3), 623-685, 05.

Cooper, Russell, John Haltiwanger and Jonathan Willis. "Search Frictions: Matching Aggregate and Establishment Observations," Journal of Monetary Economics, September 2007, 56-78.

Davis, Steven J. and John Haltiwanger. "Gross Job Creation and Destruction: Microeconomic Evidence and Macroeconomic Implications", NBER Macroeconomics Annual 1990, 123-168.

Davis, Steven J. and John Haltiwanger. "Gross Job Creation, Gross Job Destruction, and Employment Reallocation" Quarterly Journal of Economics, 1992, 819-864.

Davis, Steven J., R. Jason Faberman, and John Haltiwanger. 2012. “Labor Market Flows in the Cross Section and Over Time.” *Journal of Monetary Economics* 59(1): 1-18.

Davis, Steven J., R. Jason Faberman, John Haltiwanger, Ron Jarmin, and Javier Miranda. 2010. "Business Volatility, Job Destruction, and Unemployment." *American Economic Journal: Macroeconomics*, 2(2): 259-87.

Davis, Steven J. and John Haltiwanger. 1999. “Gross Job Flows.” Chap. 41 in *Handbook of Labor Economics* edited by Orley C. Ashenfelter and David Card, vol. 3C, pp. 2711-2805. Amsterdam, The Netherlands: Elsevier.

Davis, Steven J., John Haltiwanger, Ron Jarmin, and Javier Miranda. 2007. "Volatility and Dispersion in Business Growth Rates: Publicly Traded versus Privately Held Firms." Chap. 2 in *NBER Macroeconomics Annual 2006* edited by Daron Acemoglu, Kenneth Rogoff, and Michael Woodford, vol. 21, pp. 107-80. Cambridge, MA: MIT Press.

Davis, Steven J., John C. Haltiwanger, and Scott Schuh. 1996. *Job Creation and Destruction.* Cambridge, MA: MIT Press.

Decker, Ryan, John Haltiwanger, Ron Jarmin and Javier Miranda. 2014. “The Role of Entrepreneurship in US Job Creation and Economic Dynamism” *Journal of Economic Perspectives.*

Dunne, Timothy, Mark Roberts, and Larry Samuelson. 1989. "Plant Turnover and Gross Employment Flows in the U.S. Manufacturing Sector." *Journal of Labor Economics* 7(1): 48-71.

Ericson, Richard, and Ariel Pakes. 1995. "Markov-Perfect Industry Dynamics: A Framework for Empirical Work." *Review of Economic Studies*, 62(1): 53–82.

Fort, Teresa. 2013. "NAICS 2002 Code Assignments to the Longitudinal Business Database." Mimeo, Center for Economic Studies.

Fort, Teresa, John Haltiwanger, Ron S. Jarmin, and Javier Miranda. 2013. "How Firms Respond to Business Cycles: The Role of the Firm Age and Firm Size." *IMF Staff Papers*.

Foster, Lucia, Cheryl Grim, and John Haltiwanger. 2013. "Reallocation in the Great Recession: Cleansing or Not?" CES Working Paper no. 13-42. Washington, D.C.: Center for Economic Studies, U.S. Census Bureau.

Foster, Lucia, John Haltiwanger, and C.J. Krizan. 2001. "Aggregate Productivity Growth: Lessons from Microeconomic Evidence." Chap. 8 in *Studies in Income and Wealth* edited by Charles R. Hulten, Edwin R. Dean, and Michael J. Harper, vol. 63, *New Developments in Productivity Analysis*, pp. 303-372. Chicago, IL: National Bureau of Economic Research, University of Chicago Press.

Foster, Lucia, John Haltiwanger, and C.J. Krizan. 2006. "Market Selection, Reallocation and Restructuring in the U.S. Retail Trade Sector in the 1990s." *The Review of Economics and Statistics* 88(4): 748-58.

Foster, Lucia, John Haltiwanger, and Chad Syverson, 2008. "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" *American Economic Review*, 98(1), 394-425.

Haltiwanger, John, Ron S. Jarmin, and Javier Miranda. 2013. "Who Creates Jobs? Small vs. Large vs. Young." *Review of Economics and Statistics*.

Hecker, Daniel. 2005. "High-technology employment: a NAICS-based update," Monthly Labor Review (U.S. Dept. of Labor and U.S. Bureau of Labor Statistics), Volume 128, Number 7, July 2005: 58.

Hsieh, Chang-Tai and Peter Klenow. 2009. "Misallocation and Manufacturing TFP in China and India," Quarterly *Journal of Economics*. 124 (4): 1403–48.

Hsieh, Chang-Tai and Peter Klenow, 2014. "The Life Cycle of Plants in India and Mexico," *Quarterly Journal of Economics*.

Hugo A. Hopenhayn. 1992. "Entry, Exit, and firm Dynamics in Long Run Equilibrium." *Econometrica*, Vol. 60 (5): 1127-1150

Hopenhayn, Hugo and Richard Rogerson. 1993. "Job Turnover and Policy Evaluation: A General Equilibrium Analysis." *Journal of Political Economy* 101(5): 915-38.

Hurst, Erik and Benjamin Pugsley, 2012. "What Do Small Businesses Do?", *Brookings Papers on Economic Activity*. vol. 43(2) (Fall), pages 73-142.

Jarmin, Ron S., and Javier Miranda. 2002. "The Longitudinal Business Database." CES Working Paper no. 02-17. Washington, D.C.: Center for Economic Studies, U.S. Census Bureau.

Jovanovic, B. 1982. "Selection and the evolution of industry", in *Econometrica*, Vol. 50, No. 3, pp. 649–670.

Lazear, Edward and James Spletzer. 2012. "Hiring, Churn and the Business Cycle." *American Economic Review: Papers and Proceedings* 102(3): 575-79.

Lucas, R.E. 1978. "On the size distribution of business firms", in *The Bell Journal of Economics*, Vol. 9, No. 2, pp. 508–523.

Olley, G.Steven, and Ariel Pakes. 1996. “The Dynamics of Productivity in the Telecommunications Equipment Industry.” *Econometrica* 64(6):1263-1297.

Reedy, E.J. and Robert E. Litan. 2011. “Starting Smaller; Staying Smaller: America's Slow Leak in Job Creation.” Kauffman Foundation Research Series: Firm Foundation and Economic Growth. At http://www.kauffman.org/research-and-policy/starting-smaller-staying-smaller-americas-slow-leak-in-job-creation.aspx.

Restuccia, Diego and Richard Rogerson. 2008. “Policy Distortions and Aggregate Productivity with Heterogeneous Plants,” Review of Economic Dynamics, 11(October): 707-720.

Schoar, Antoinette. 2010. "The Divide between Subsistence and Transformational Entrepreneurship" in Joshua Lerner and Scott Stern (Eds.) *NBER Innovation Policy and the Economy*.

Syverson, Chad. 2011. “What Determines Productivity?” *Journal of Economic Literature* 49(2): 326-65.

Törnqvist, Leo, Pentti Vartia, and Yrjö O. Vartia. 1985. “How Should Relative Changes be Measured?” *The American Statistician* 39(1): 43-6.

Wooldridge, Jeffrey. 2002. “Inverse Probability Weighted M-estimators for Sample Selection, Attrition, and Stratification.” *Portuguese Economic Journal*, 1, 117:139.

Wooldridge, Jeffrey. 2006. “Inverse Probability Weighted Estimation for General Missing Data Problems.” Centre for Microdata Methods and Practice. No. CWP05/04.

Table 1: The Share of Revenue and Job Creation Accounted for by High Growth Firms

	High Growth Firms	
	Revenue	Employment
Share of Gross Creation	49.3%	58.0%
Share of Firms	12.5%	16.5%

Note: Tabulations from Revenue enhanced LBD subset 1998-2000, 2003-2011.

Table 2a: VAR relating Net Employment and Net Revenue Growth

	Dependent Variable	
Explanatory Variables:	Net Employment Growth	Net Revenue Growth
Lagged Net Employment Growth	-0.163	0.039
	(0.000)	(0.000)
Lagged Net Revenue Growth	0.187	0.018
	(0.000)	(0.000)

Note: Estimated specifications using Revenue enhanced LBD subset 1998-2000, 2003-2011. Standard errors in parentheses. Specifications are weighted by employment and revenue, respectively.

Table 2b: VAR relating probability of being High Growth Revenue and High Growth Employment Firms

	Dependent Variable	
Explanatory Variables:	High Growth Employment Firms	High Growth Revenue Firms
Lagged High Growth Employment Firms	0.056	0.061
	(0.000)	(0.000)
Lagged High Growth Revenue Firms	0.112	0.123
	(0.000)	(0.000)

Note: Estimated specifications Revenue enhanced LBD subset 1998-2000, 2003-2011. Standard errors in parentheses. Specifications are weighted by employment and revenue, respectively.

Table 3: Adjusted R-squared for Effects Accounting for High Growth Firms

	High Growth Revenue Firms	High Growth Employment Firms
Industry	0.067	0.041
Age	0.019	0.051
Base Year Size	0.009	0.032
Average Size	0.003	0.014
Year	0.010	0.004
State	0.008	0.002
All Effects (Base Year Size)	0.096	0.090
All Effects (Average Size)	0.093	0.079

Note: Estimated adjusted R-squared with dependent variable using the Revenue enhanced LBD subset 1998-2000, 2003-2011. Specifications are weighted by revenue (column 1) and employment respectively.

Table 4: The Role of Industry Groupings in Accounting for High Growth Firms

Dependent Variable: High Growth Firm Industry Effects		
	Revenue	Employment
Explanatory Variables:		
Tradable	-0.011	-0.052
	(0.009)	(0.007)
Construction	0.020	0.027
	(0.014)	(0.011)
High Tech	0.020	0.033
	(0.018)	(0.013)
Bio Tech	-0.015	-0.012
	(0.038)	(0.028)
Energy	0.077	-0.005
	(0.016)	(0.012)

Note: Dependent variable are the estimated industry effects on high growth firms.

Table 5: Correlations of High Growth Industry Effects with Summary Measures of First, Second and Third Moments of Industry Distributions

	Rev (GR)	Rev (HG)	Emp (GR)	Emp (HG)	Rev (HD)	Emp (HD)	Rev (90-10)	Emp (90-10)	Rev (Skew)	Emp (Skew)
Rev (GR)	1	0.50	0.38	-0.04	-0.45	-0.26	-0.06	-0.17	0.13	0.18
Rev (HG)		1.00	0.16	0.39	0.41	0.33	0.74	0.39	-0.05	0.08
Emp (GR)			1.00	0.49	-0.15	0.02	0.01	0.26	0.22	0.48
Emp (HG)				1.00	0.43	0.83	0.54	0.94	0.01	0.27
Rev (HD)					1.00	0.56	0.81	0.53	-0.38	-0.02
Emp (HD)						1.00	0.59	0.94	-0.12	-0.09
Rev (90-10)							1.00	0.62	-0.26	0.01
Emp (90-10)								1.00	-0.08	0.09
Rev (Skew)									1.00	0.13
Emp (Skew)										1.00

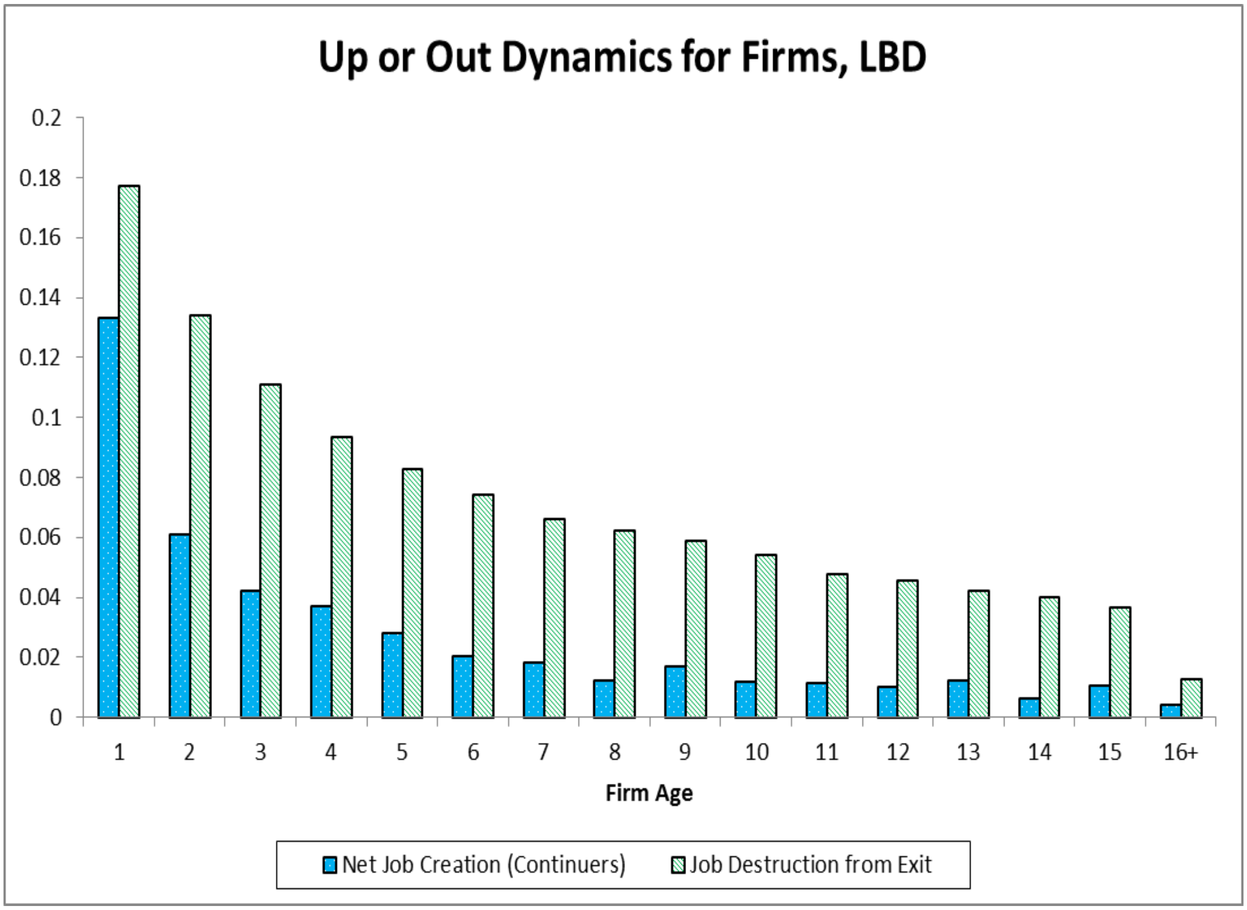
Note: Rev=Revenue, Emp=Employment, GR= net growth, HG=high growth industry effect, HD=high decline industry effect, 90-10=activity weighted 90-10 differential (employment weights for Emp and revenue weights for Rev). Skew=(90-50)-(50-10) (activity weighted).

Table 6: Correlations of High Growth State Effects with Summary Measures of First, Second and Third Moments of State Distributions

	Rev (GR)	Rev (HG)	Emp (GR)	Emp (HG)	Rev (HD)	Emp (HD)	Rev (90-10)	Emp (90-10)	Rev (Skew)	Emp (Skew)
Rev (GR)	1	0.72	0.55	0.30	-0.19	0.09	0.30	0.24	0.44	0.15
Rev (HG)		1.00	0.54	0.67	0.40	0.52	0.81	0.62	0.30	0.33
Emp (GR)			1.00	0.50	0.03	0.11	0.28	0.34	0.22	0.46
Emp (HG)				1.00	0.60	0.89	0.77	0.97	-0.05	0.39
Rev (HD)					1.00	0.64	0.81	0.62	-0.55	0.26
Emp (HD)						1.00	0.72	0.95	-0.13	0.08
Rev (90-10)							1.00	0.76	-0.15	0.36
Emp (90-10)								1.00	-0.09	0.26
Rev (Skew)									1	-0.03
Emp (Skew)										1.00

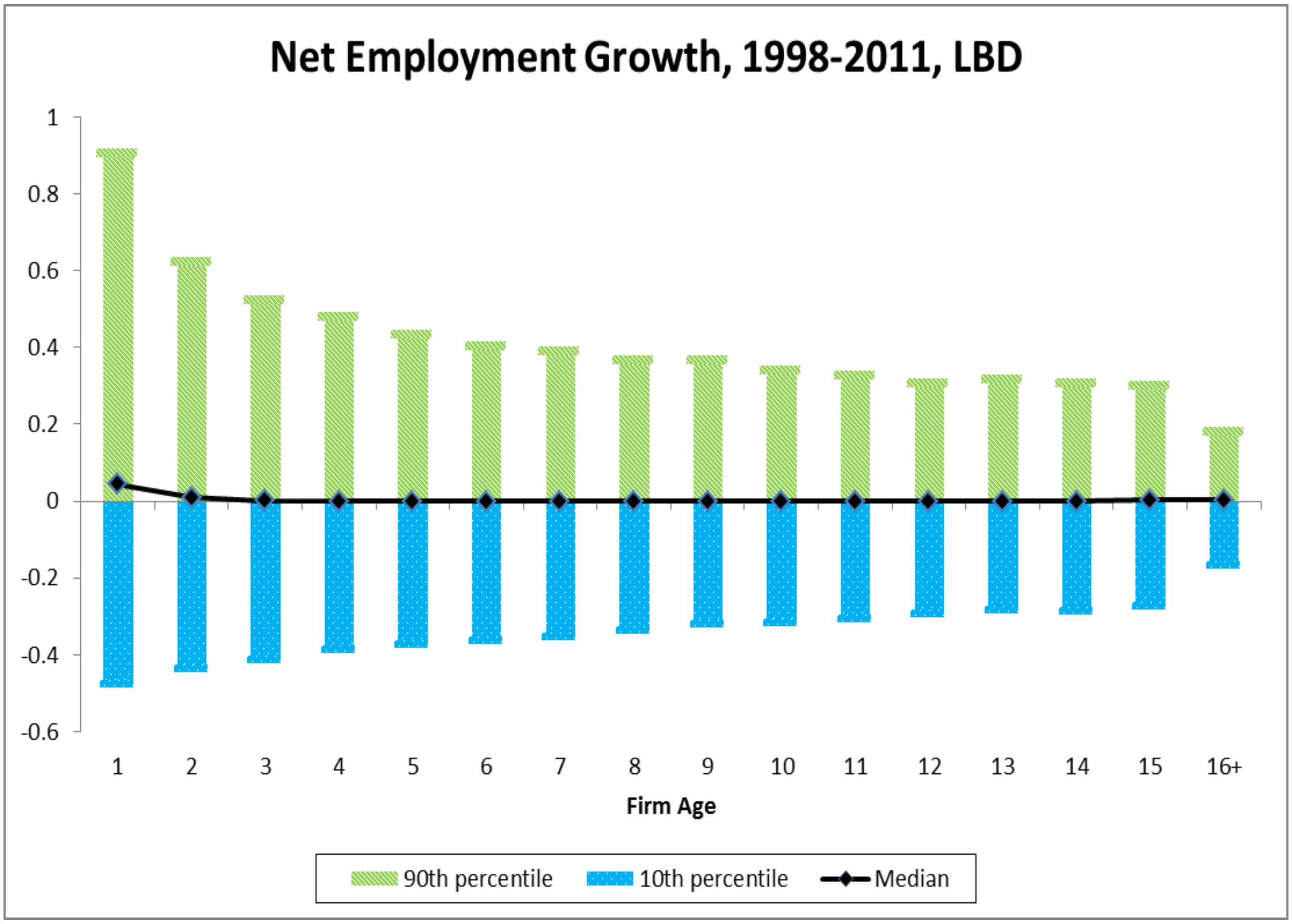
Note: Rev=Revenue, Emp=Employment, GR= net growth, HG=high growth industry effect, HD=high decline industry effect, 90-10=activity weighted 90-10 differential (employment weights for Emp and revenue weights for Rev). Skew=(90-50)-(50-10) (activity weighted).

Figure 1



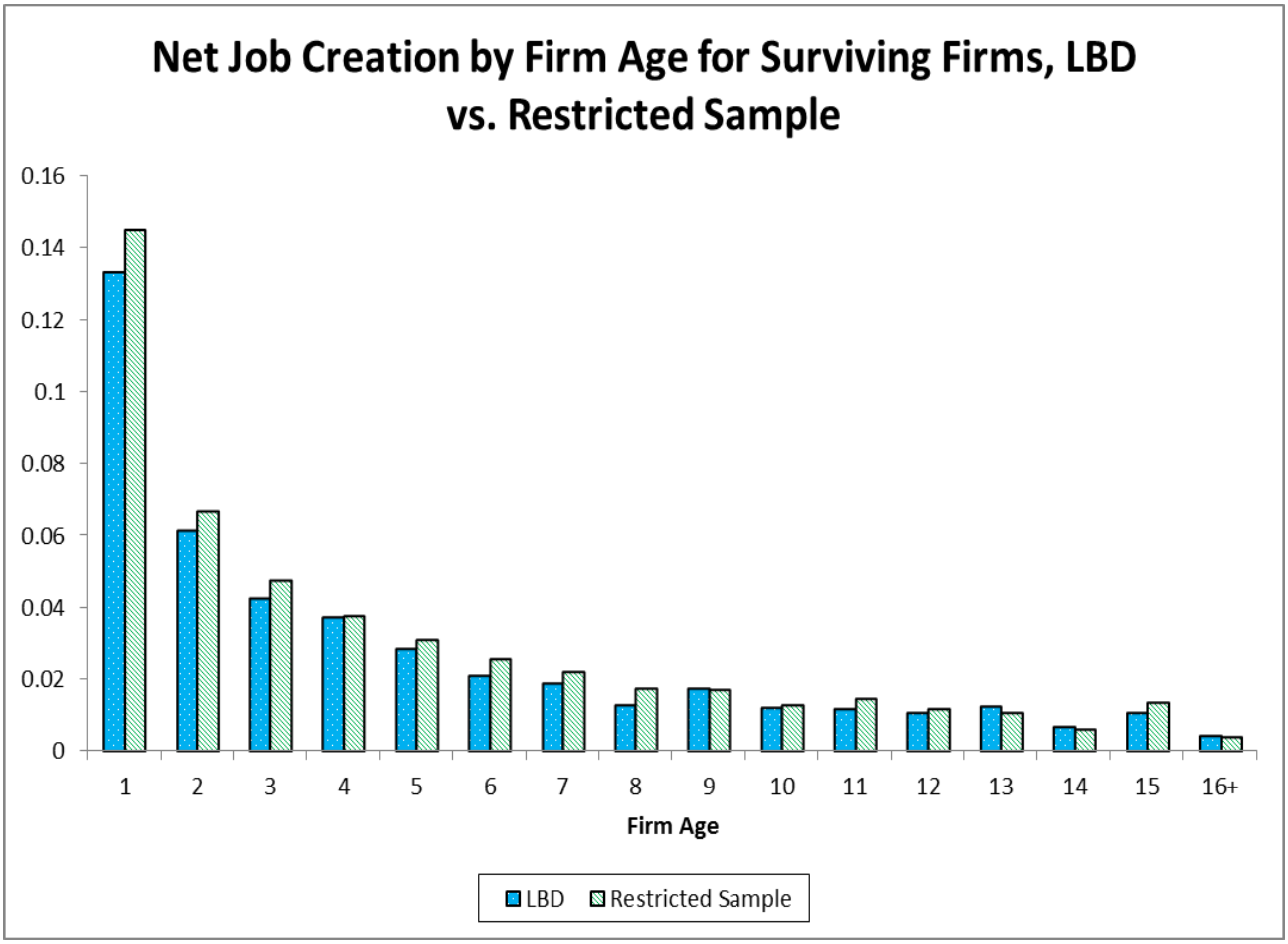
Source: Statistics computed from the Longitudinal Business Database 1998-2000, 2003-2011.
Notes: Figure 1 shows patterns of net employment growth for continuing firms and job destruction from firm exit for firms age 1 and older.

Figure 2



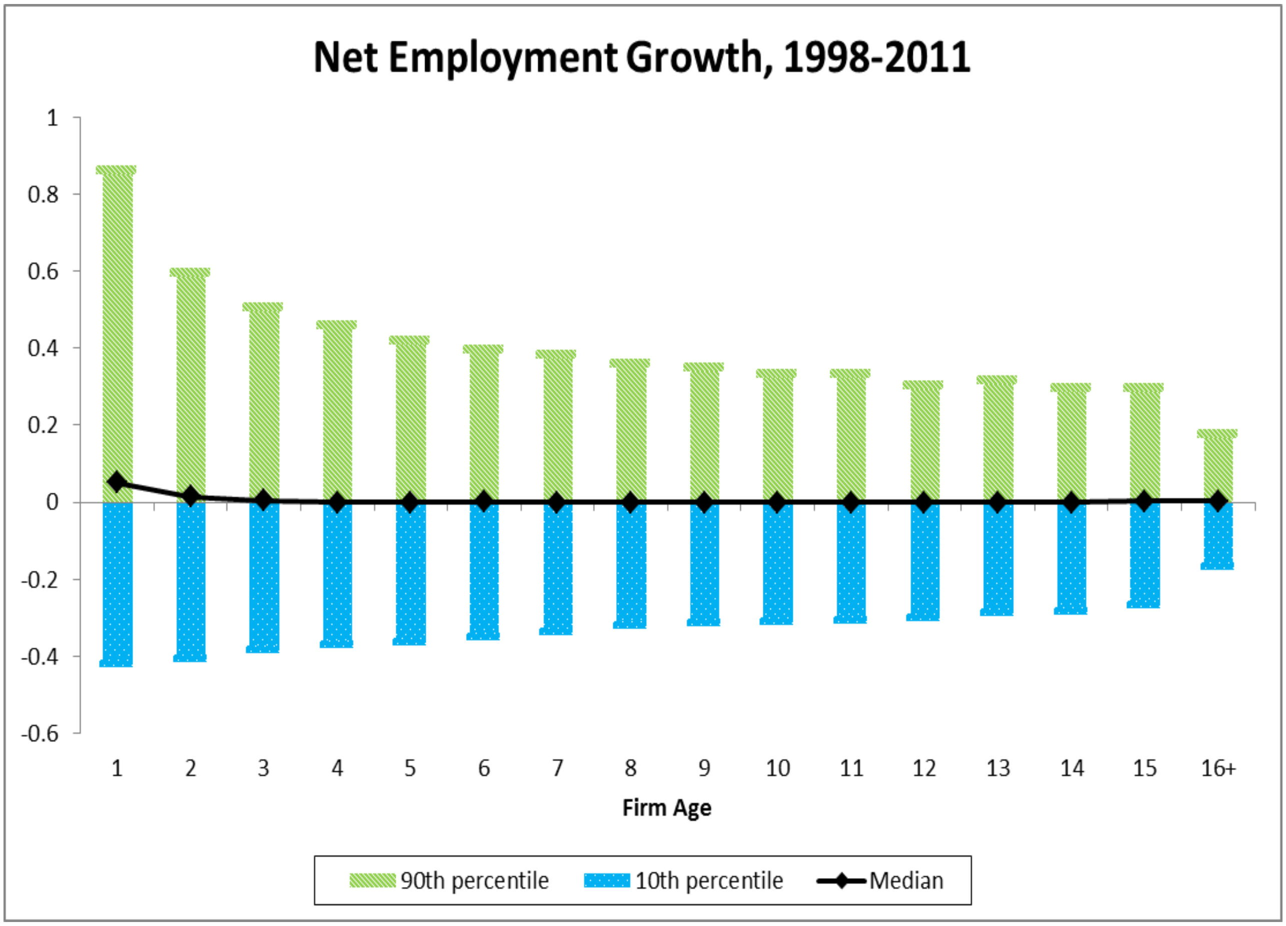
Source: Statistics computed from the Longitudinal Business Database 1998-2000, 2003-2011.
Notes: The 90th, 10th, and median are all based on the employment-weighted firm level growth rate distribution for each firm.

Figure 3



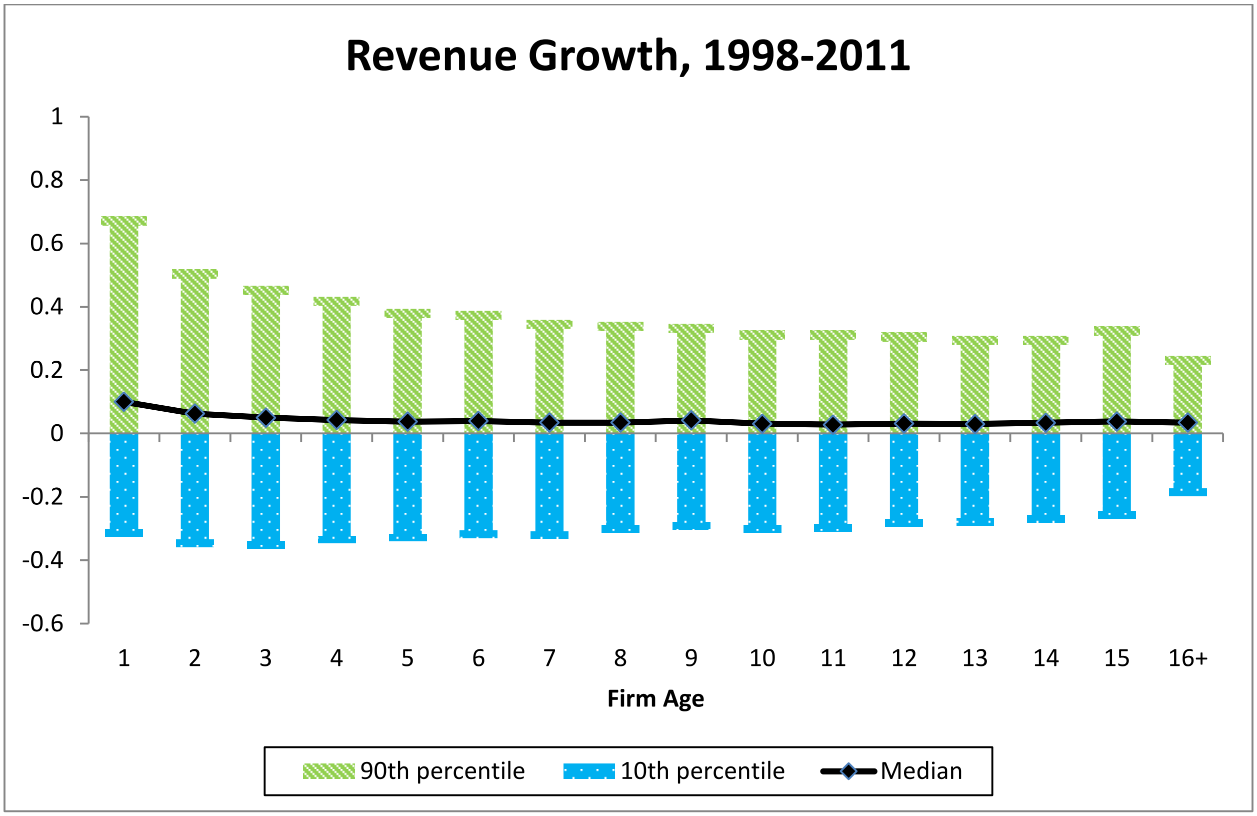
Source: Statistics computed from the Longitudinal Business Database 1998-2000, 2003-2011 and Revenue enhanced LBD subset 1998-2000, 2003-2011.
Notes: Figure 3 compares net job creation rate by firm age for surviving firms in the LBD versus the restricted sample.

Figure 4



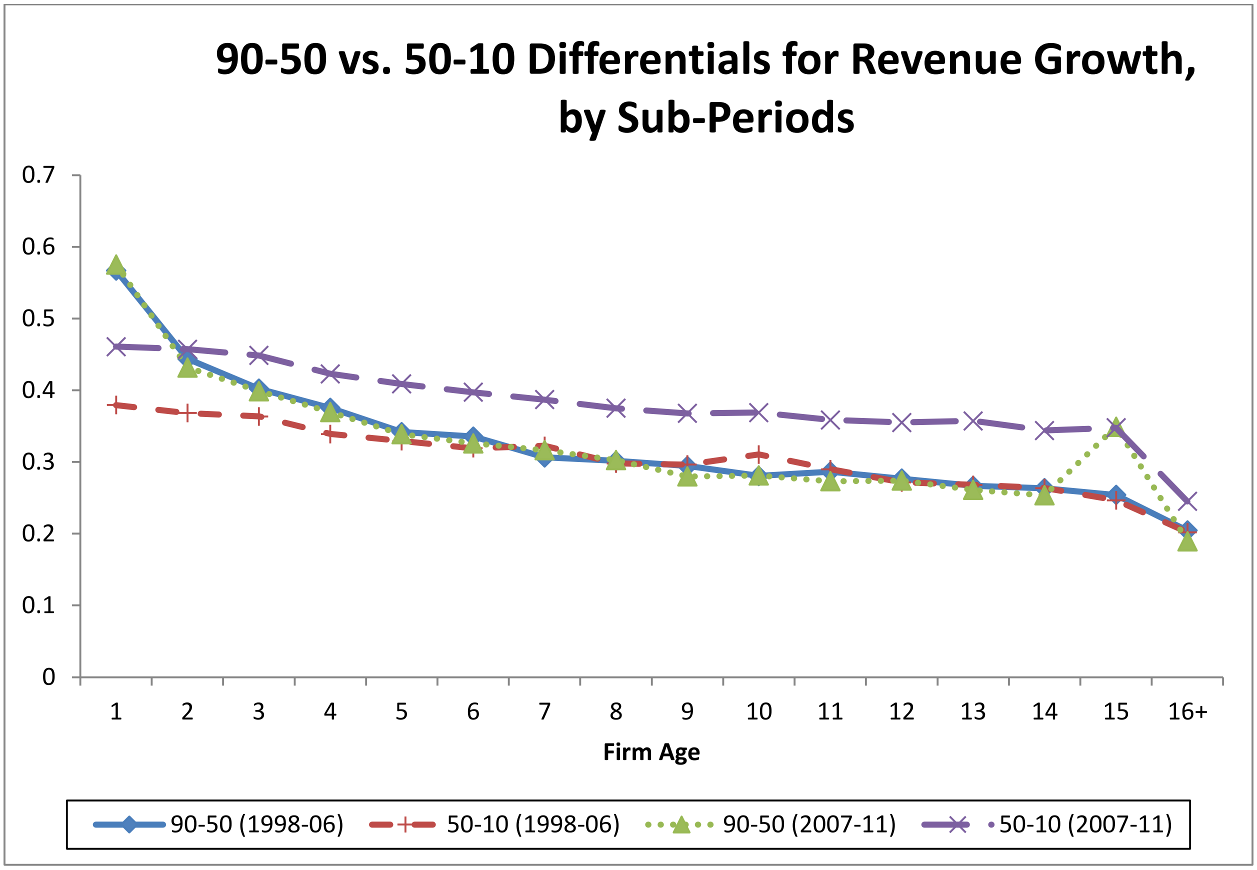
Source: Statistics computed from the Revenue enhanced LBD subset 1998-2000, 2003-2011.
Notes: The 90th, 10th, and median are all based on the employment-weighted firm level employment growth rate distribution for each firm.

Figure 5



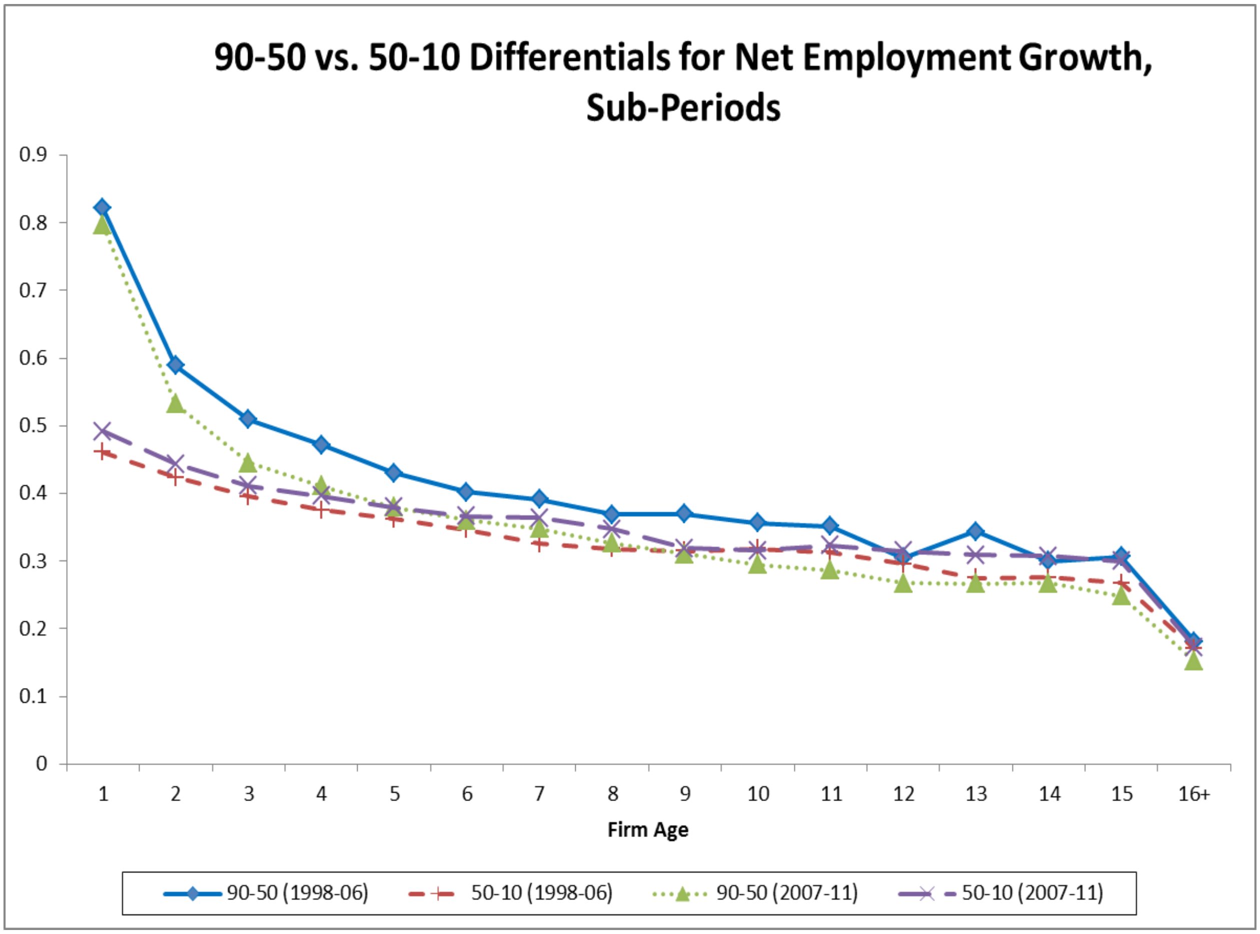
Source: Statistics computed from the Revenue enhanced LBD subset 1998-2000, 2003-2011.
Notes: The 90th, 10th, and median are all based on the revenue-weighted firm level revenue growth rate distribution for each firm.

Figure 6a



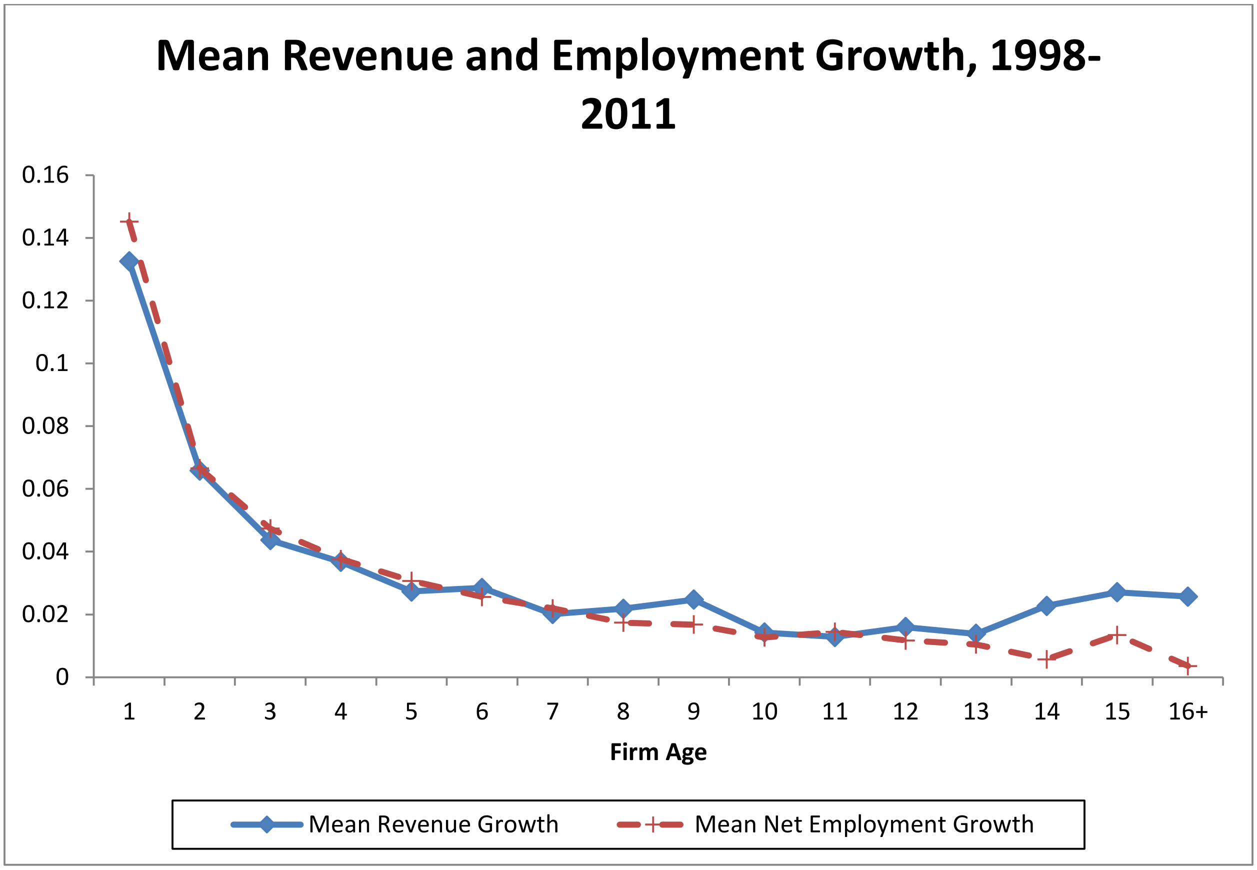
Source: Statistics computed from the Revenue enhanced LBD subset 1998-2000, 2003-2011.
Notes: The 90th, 10th, and median are all based on the revenue-weighted firm level revenue growth rate distribution for each firm.

Figure 6b



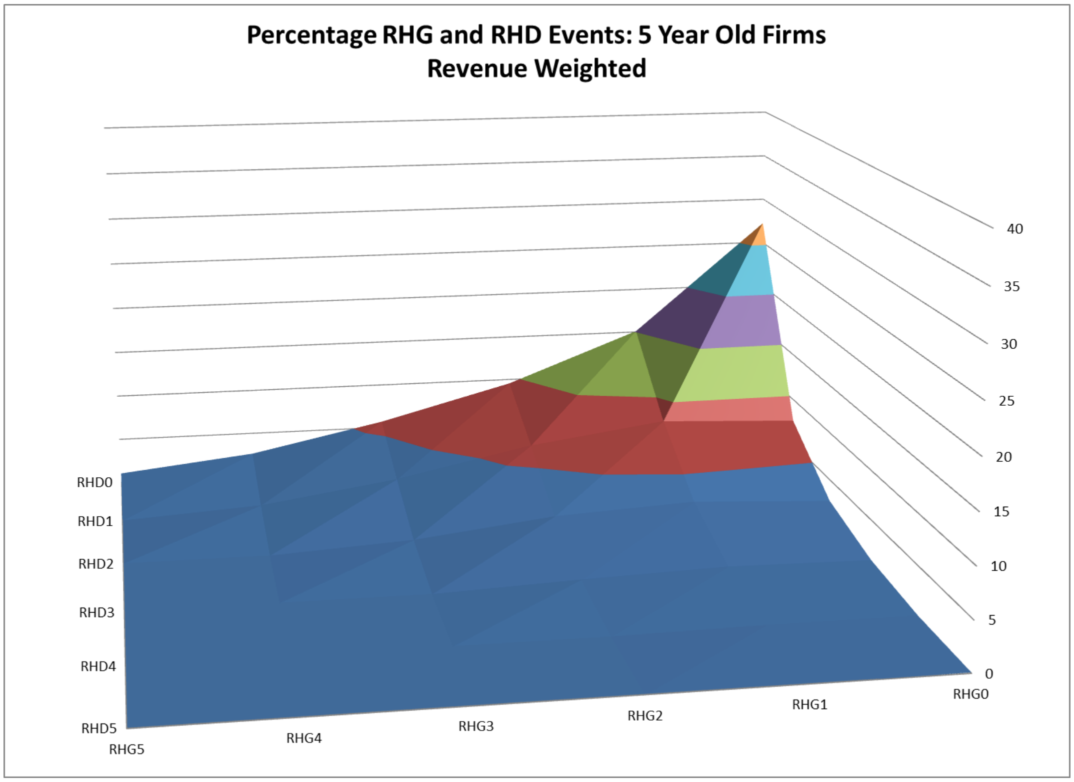
Source: Statistics computed from the Revenue enhanced LBD subset 1998-2000, 2003-2011.
Notes: The 90th, 10th, and median are all based on the employment-weighted firm level employment growth rate distribution for each firm.

Figure 6c:



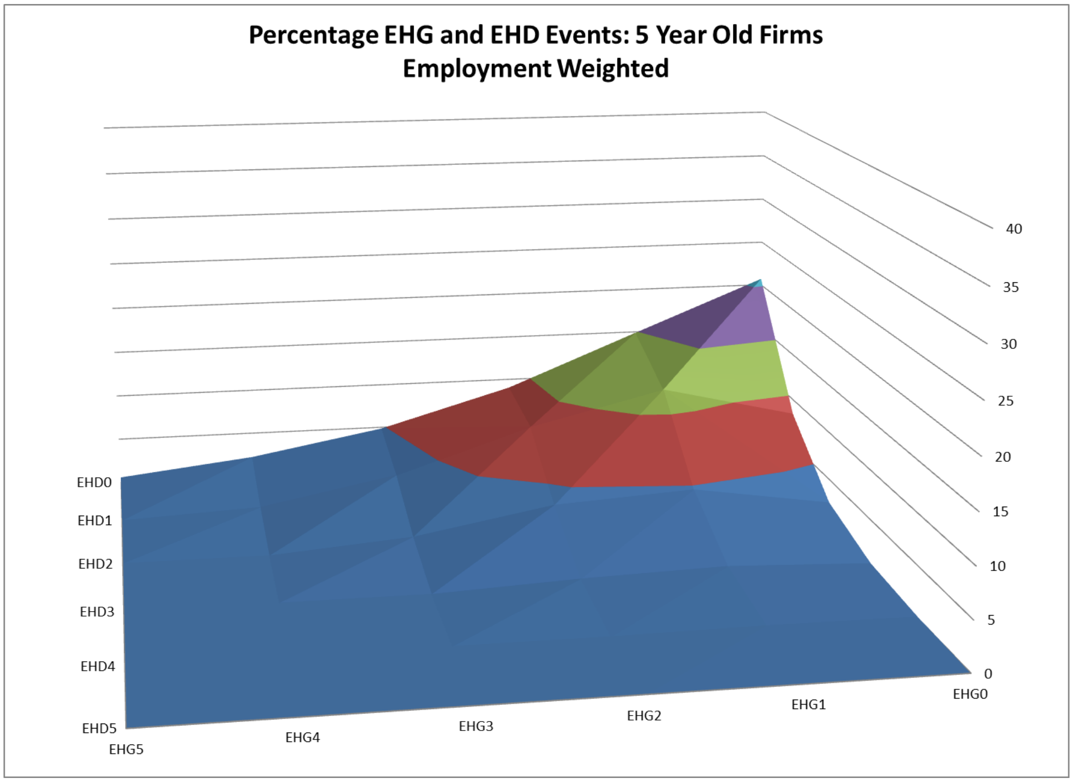
Source: Statistics computed from the Revenue enhanced LBD subset 1998-2000, 2003-2011.
Notes: The mean net employment growth is the employment-weighted average firm level employment growth rate for each firm age. The mean revenue growth is the revenue-weighted average firm level revenue growth rate for each firm age.

Figure 7a:



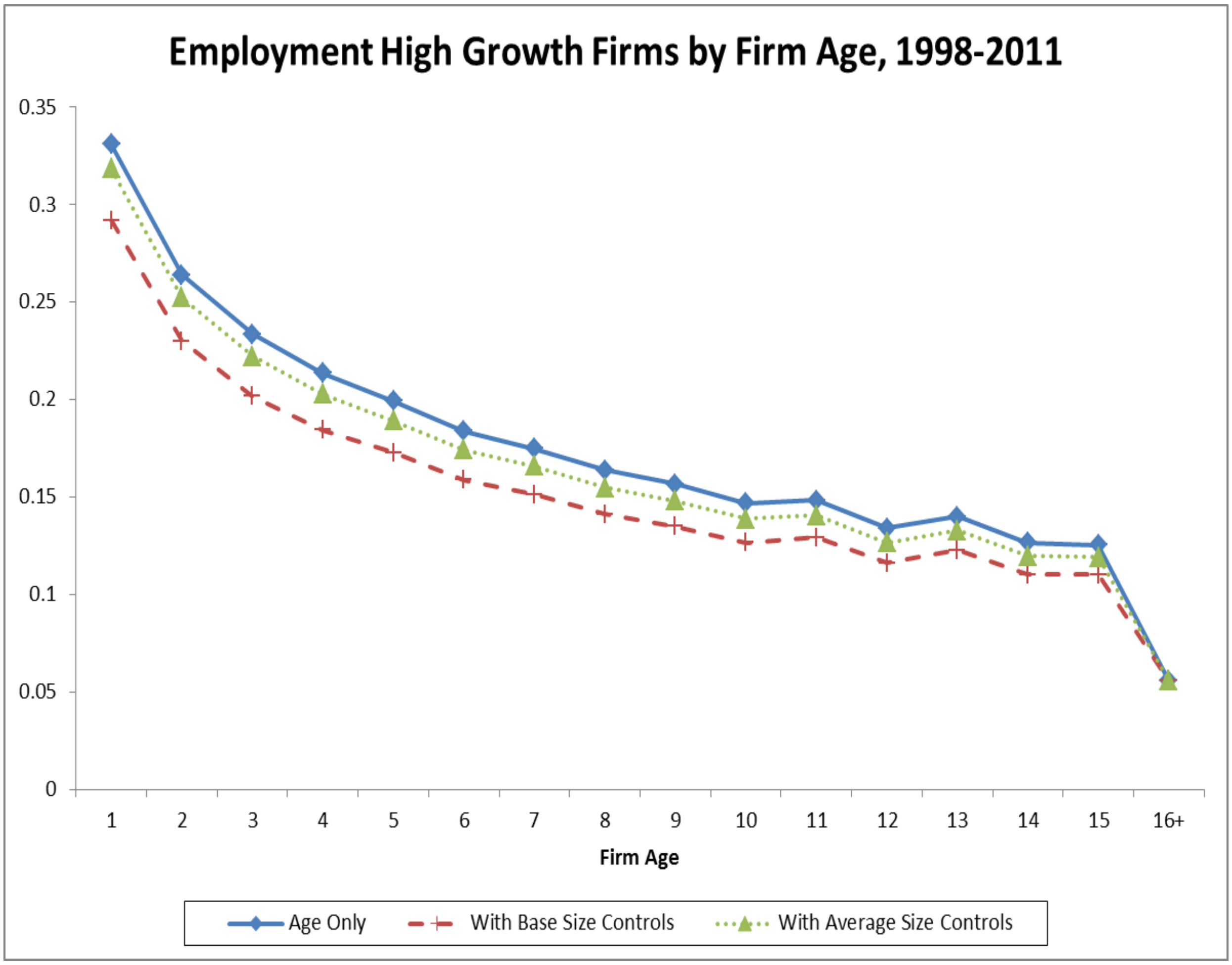
Source: Statistics computed from the Revenue enhanced LBD subset 1998-2000, 2003-2011. RHG = revenue high growth. RHD = revenue high decline. Reported shares are revenue weighted.

Figure 7b:



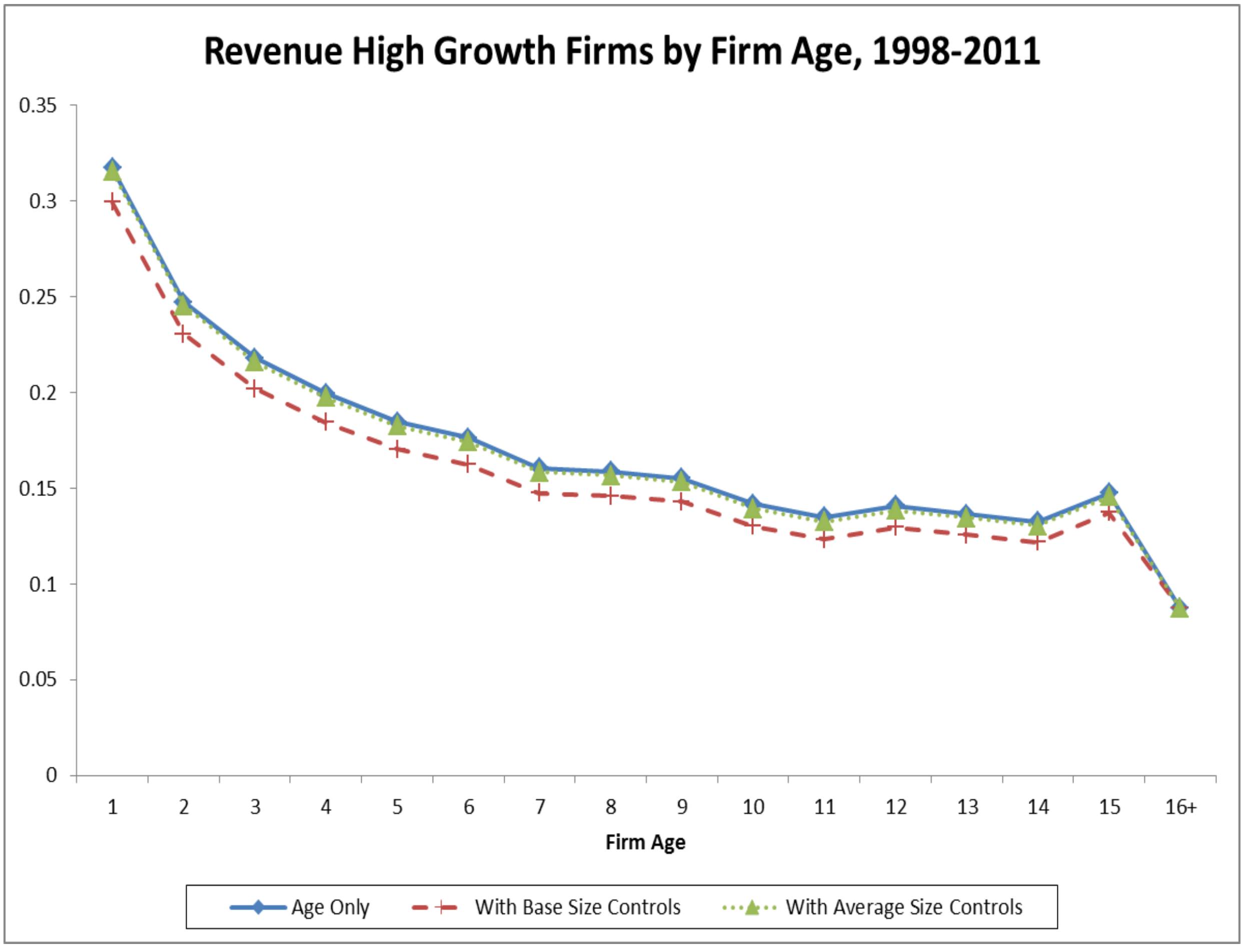
Source: Statistics computed from the Revenue enhanced LBD subset 1998-2000, 2003-2011. EHG = employment high growth. RHD = employment high decline. Reported shares are employment weighted.

Figure 8:



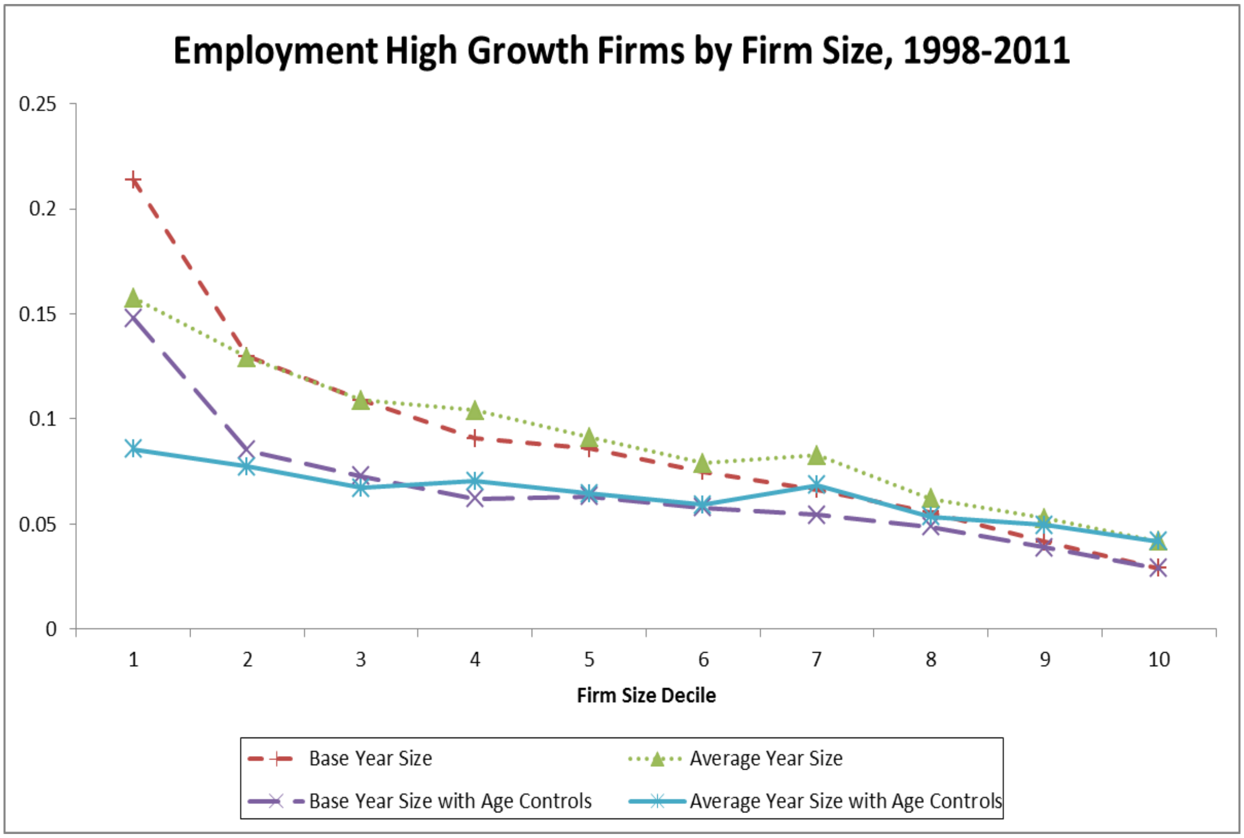
Source: Statistics computed from the Revenue enhanced LBD subset 1998-2000, 2003-2011. Reported are estimated effects of linear probability models on controls as listed. All coefficients are reported relative to unconditional mean for 16+.

Figure 9:



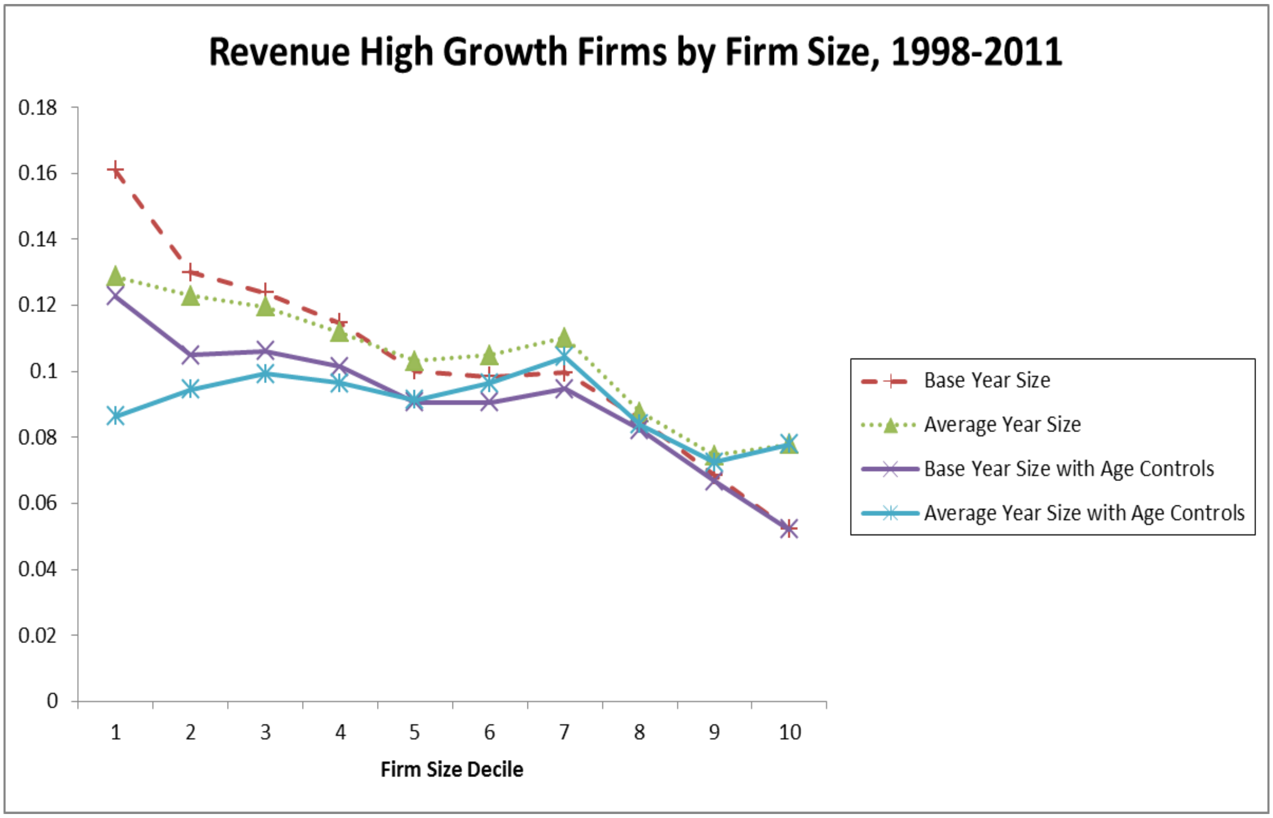
Source: Statistics computed from the Revenue enhanced LBD subset 1998-2000, 2003-2011. Reported are estimated effects of linear probability models on controls as listed. All coefficients are reported relative to unconditional mean for 16+.

Figure 10:



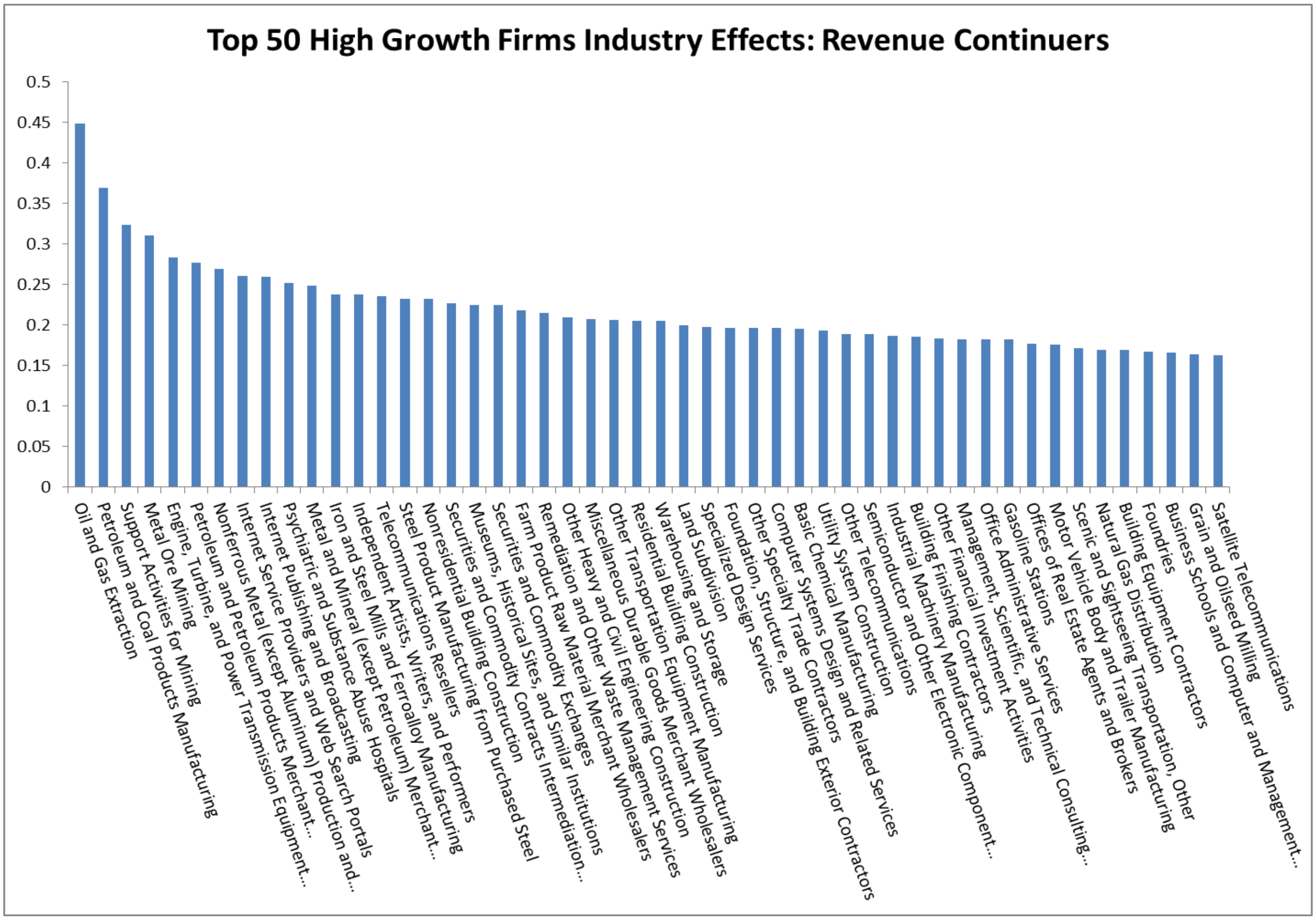
Source: Statistics computed from the Revenue enhanced LBD subset 1998-2000, 2003-2011. Reported are estimated effects of linear probability models on controls as listed. All coefficients are reported relative to unconditional mean for 16+.

Figure 11:



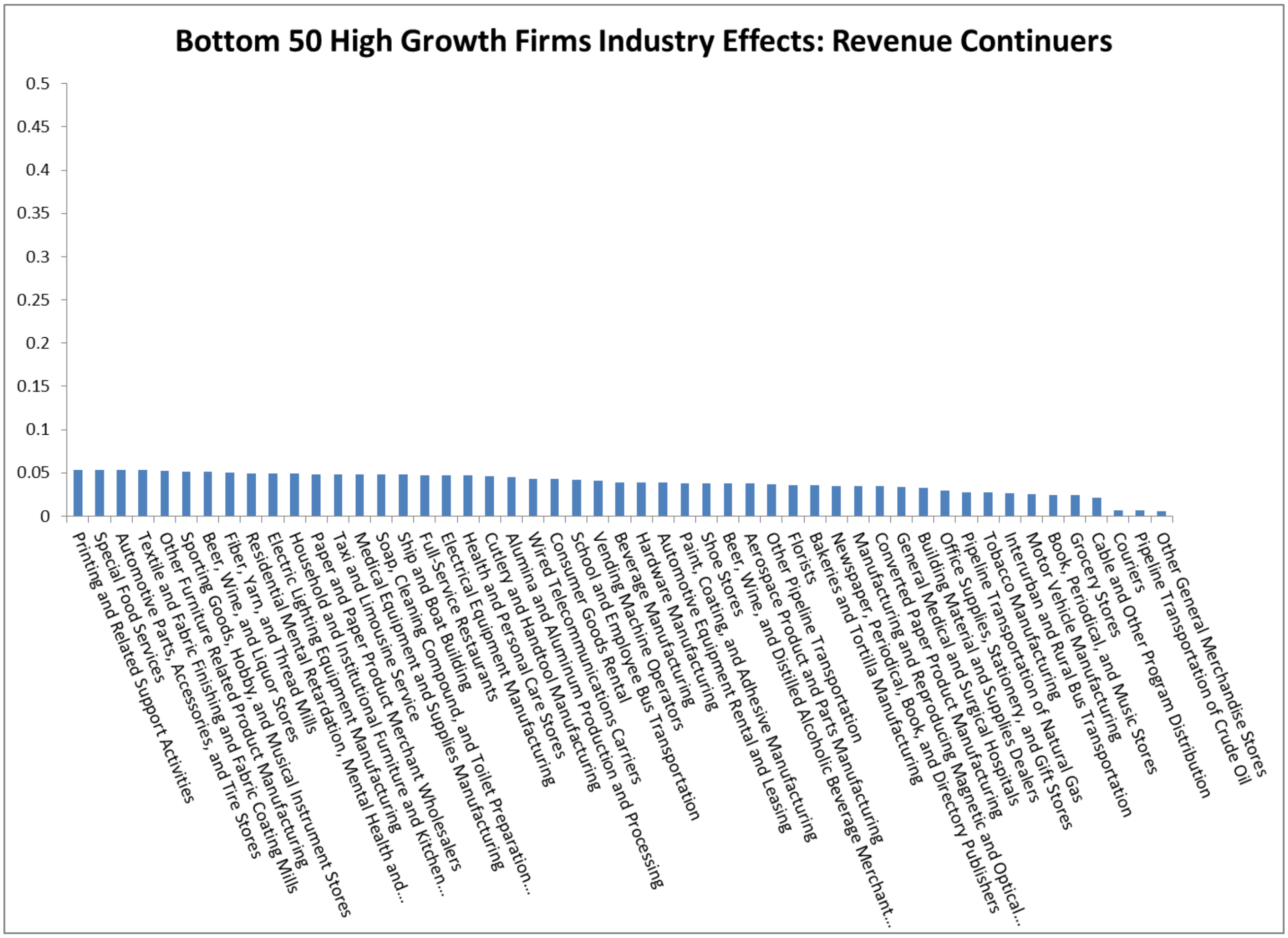
Source: Statistics computed from the Revenue enhanced LBD subset 1998-2000, 2003-2011. Reported are estimated effects of linear probability models on controls as listed. All coefficients are reported relative to unconditional mean for 16+.

Figure 12a:



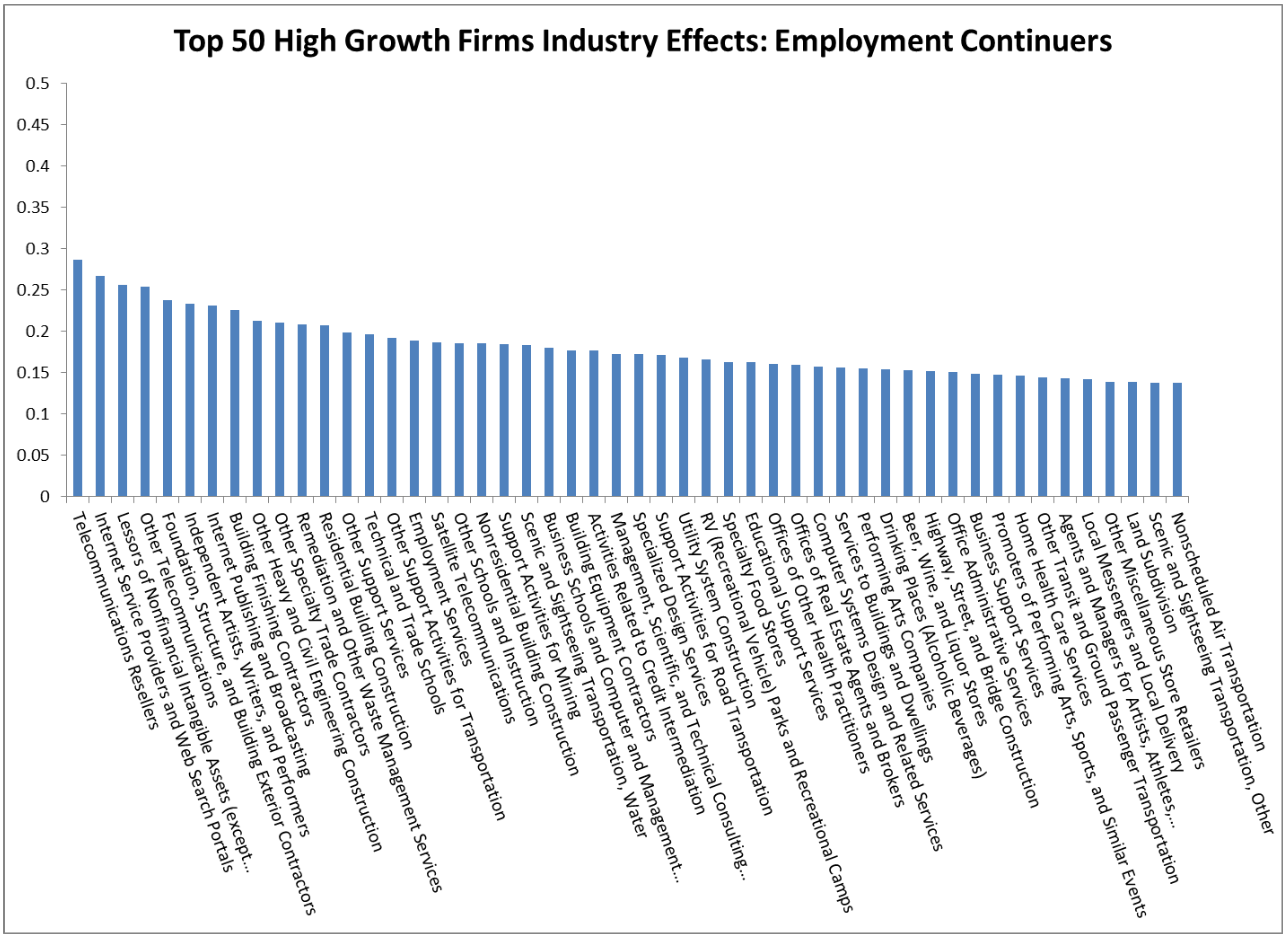
Source: Statistics computed from the Revenue enhanced LBD subset 1998-2000, 2003-2011. Reported are estimated effects of linear probability models on industry effects.

Figure 12b:



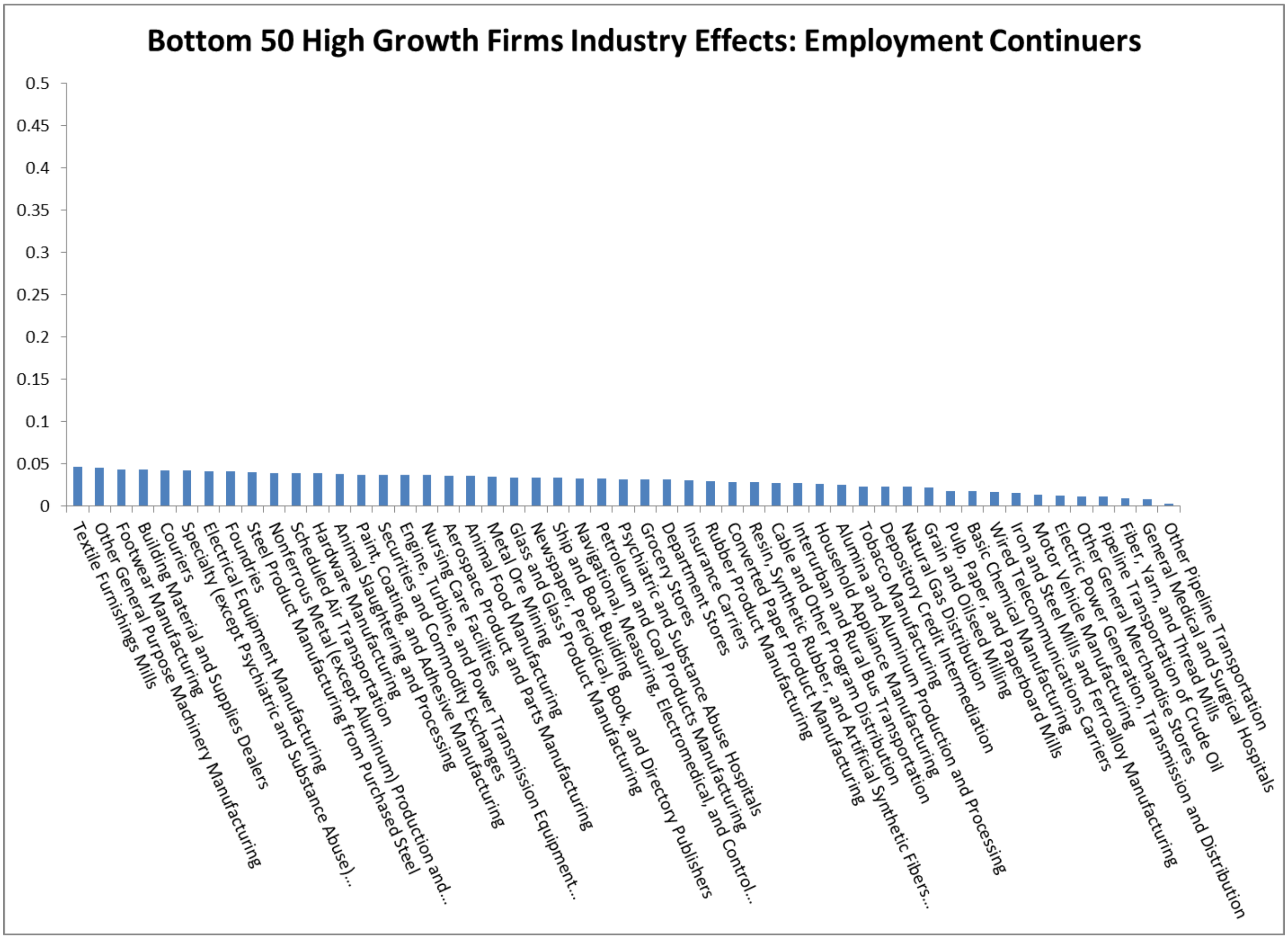
Source: Statistics computed from the Revenue enhanced LBD subset 1998-2000, 2003-2011. Reported are estimated effects of linear probability models on industry effects.

Figure 13a:



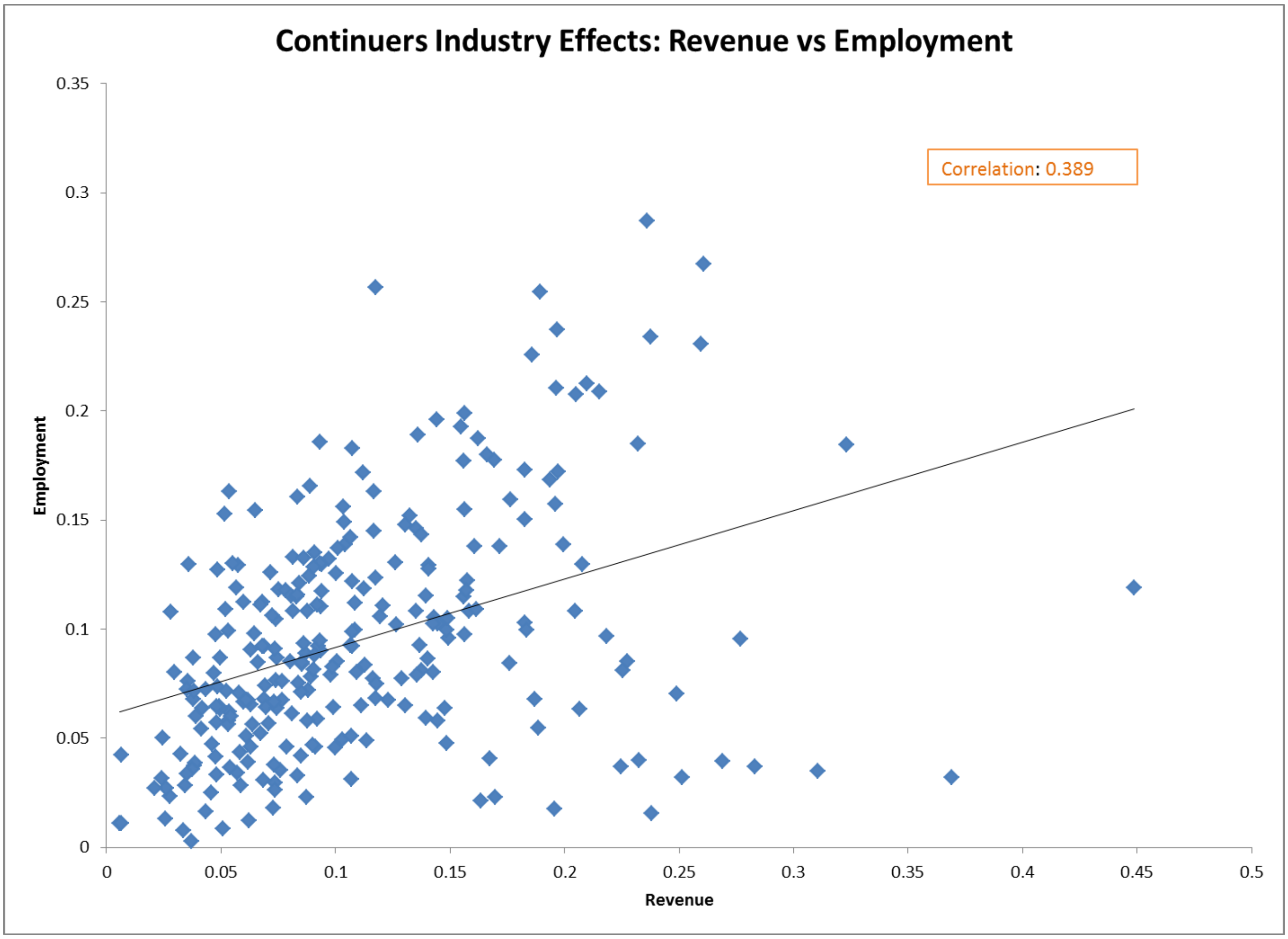
Source: Statistics computed from the Revenue enhanced LBD subset 1998-2000, 2003-2011. Reported are estimated effects of linear probability models on industry effects.

Figure 13b:



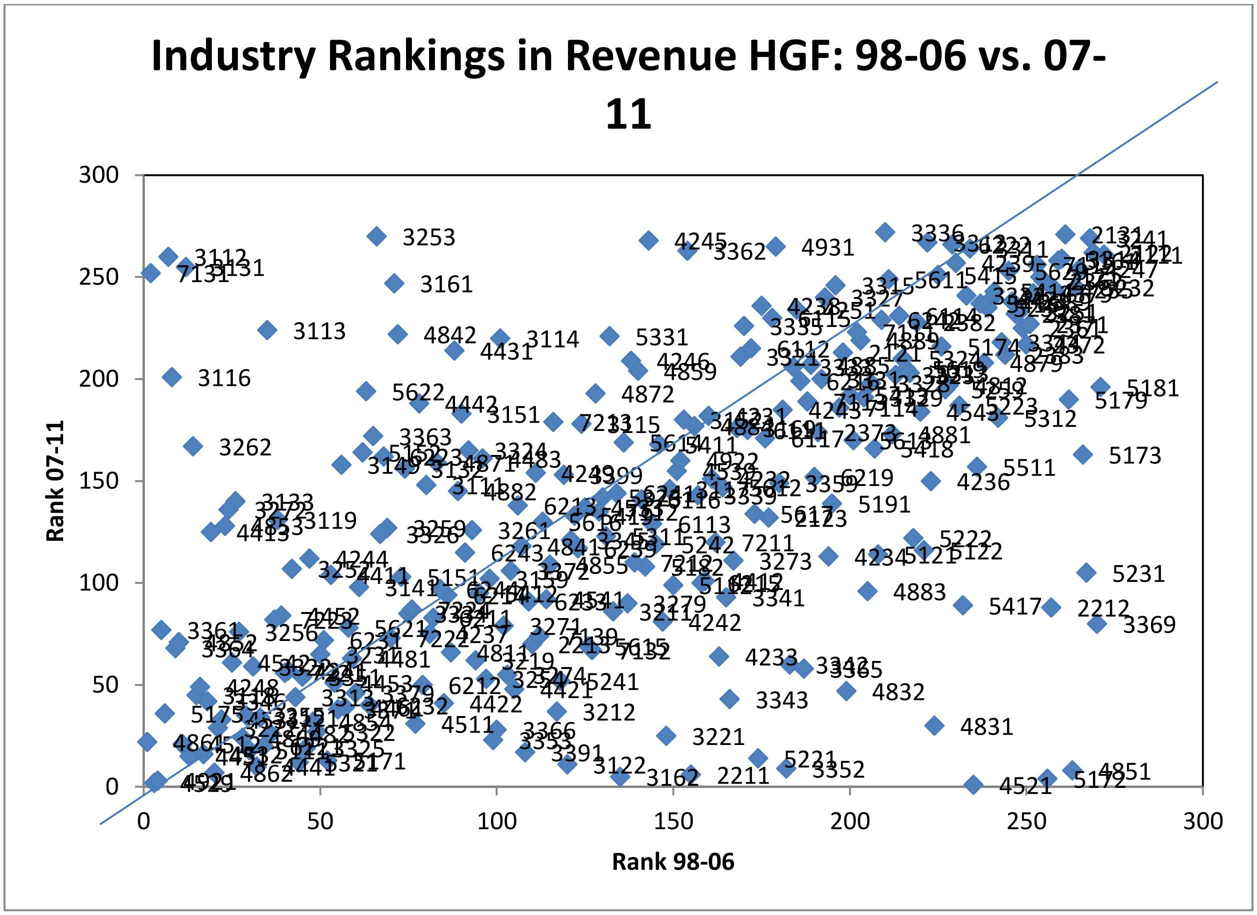
Source: Statistics computed from the Revenue enhanced LBD subset 1998-2000, 2003-2011. Reported are estimated effects of linear probability models on industry effects.

Figure 14:



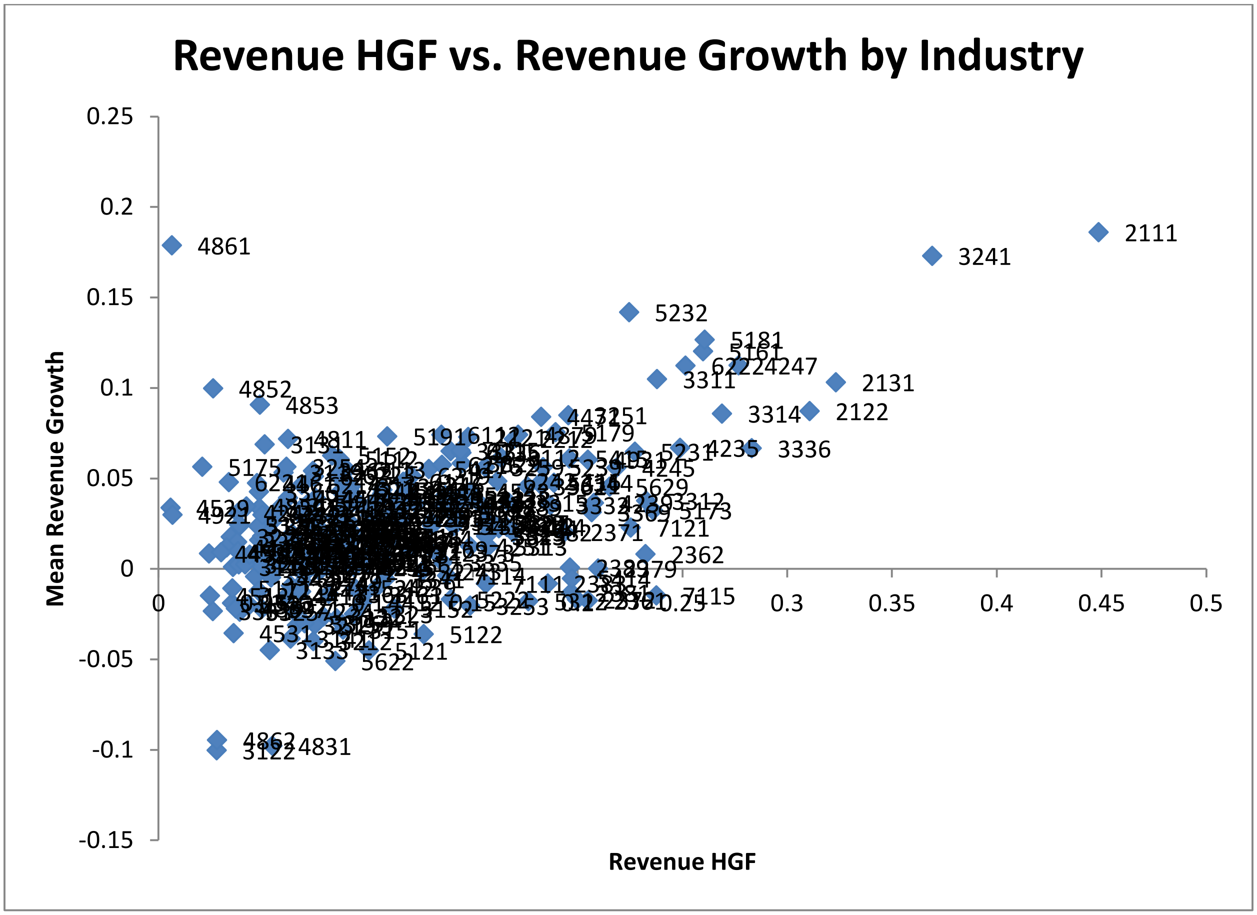
Source: Statistics computed from the Revenue enhanced LBD subset 1998-2000, 2003-2011. Reported are estimated effects of linear probability models on industry effects.

Figure 15:



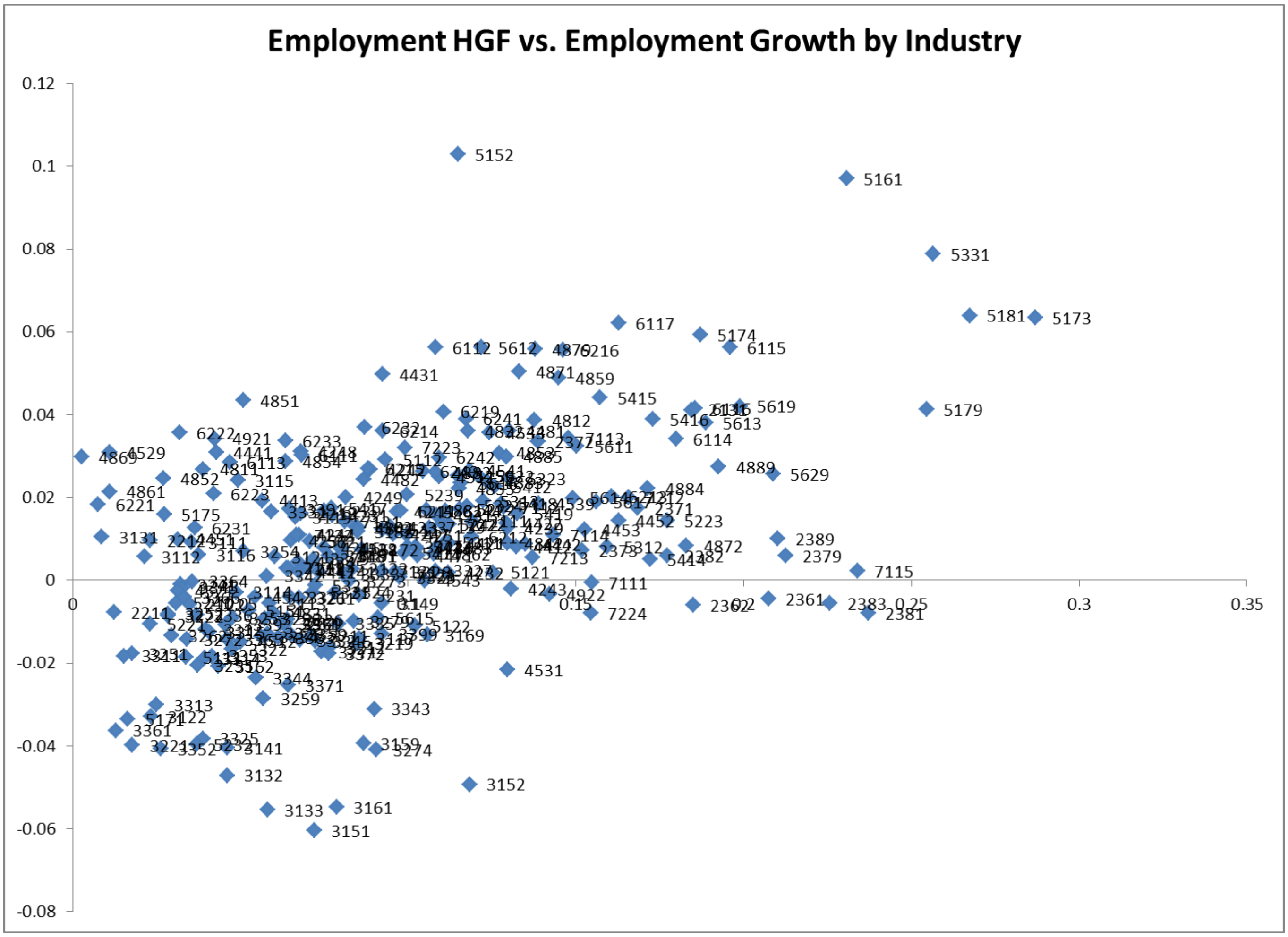
Source: Statistics computed from the Revenue enhanced LBD subset 1998-2000 and 2003-2011. The rankings for 1998-06 use the estimates from the 1998-06 period (except for 2001 and 2002) and the rankings of 2007-11 use the estimates from the 2007-11 period. Reported are estimated effects of linear probability models on industry effects.

Figure 16a:



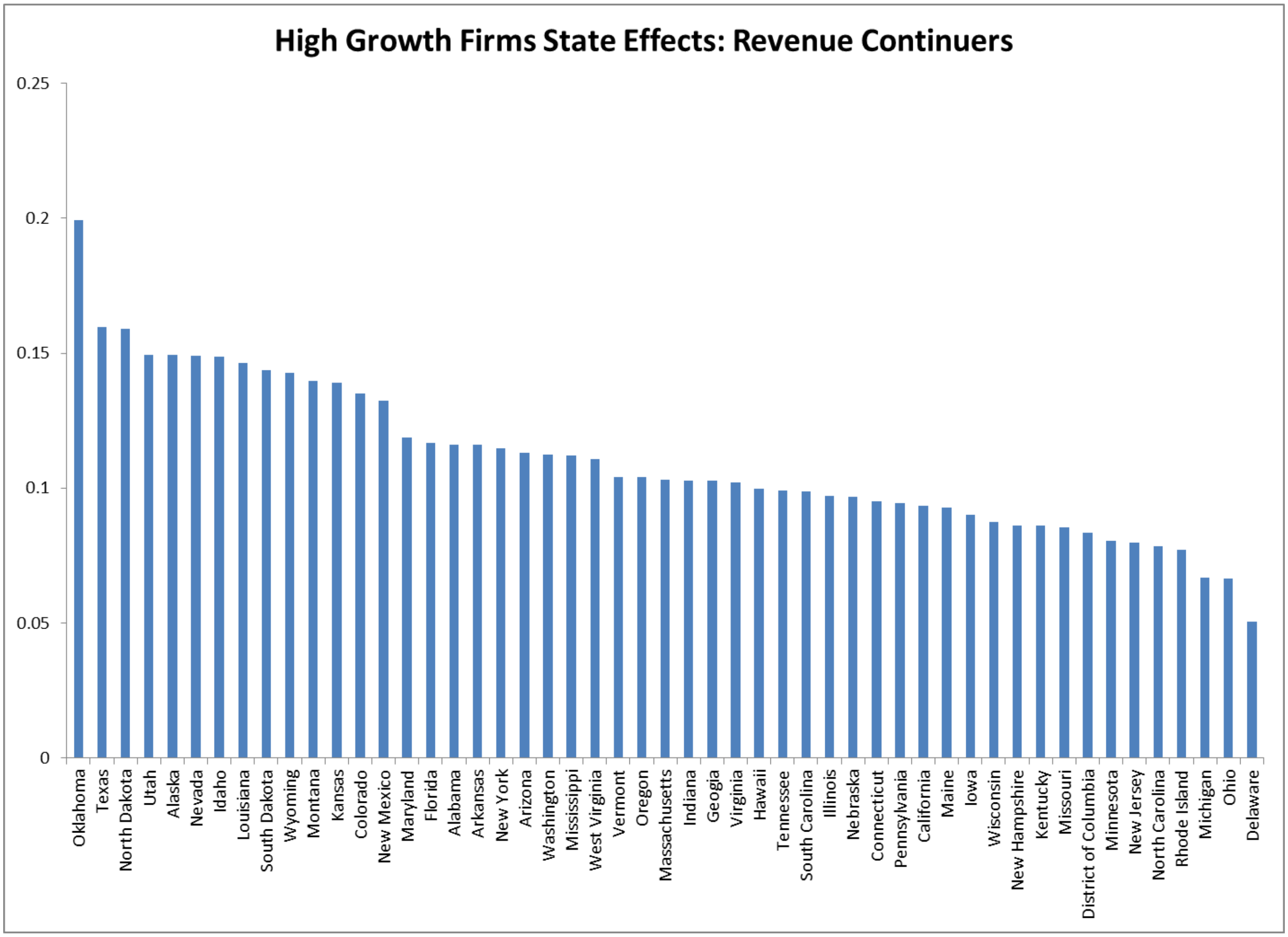
Source: Statistics computed from the Revenue enhanced LBD subset 1998-2000, 2003-2011. Reported Revenue HGF are estimated effects of linear probability models on industry effects. Mean Revenue Growth is revenue weighted mean revenue growth for firms.

Figure 16b:



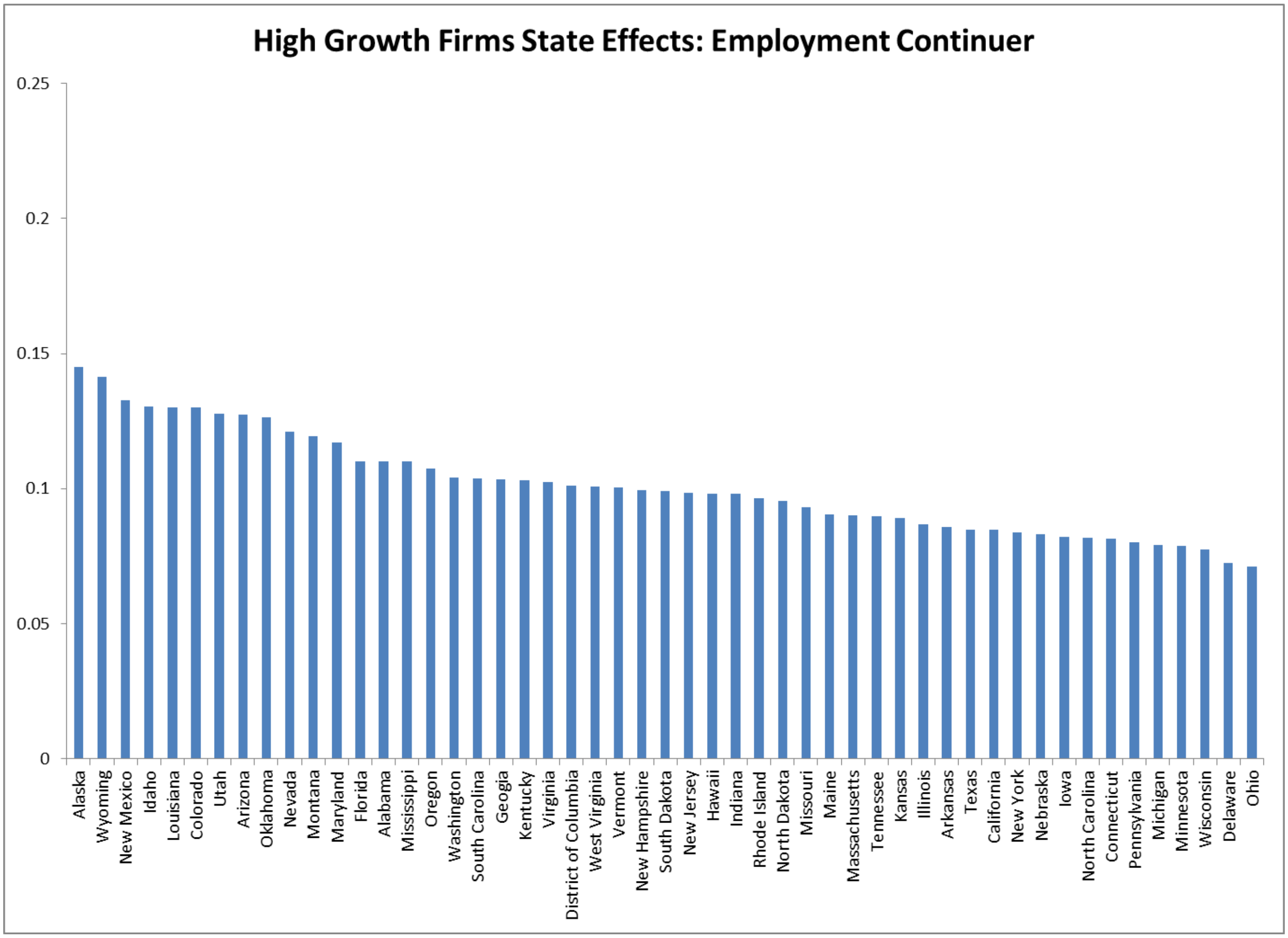
Source: Statistics computed from the Revenue enhanced LBD subset 1998-2000, 2003-2011. Reported Employment HGF are estimated effects of linear probability models on industry effects. Mean Employment growth is employment weighted mean employment growth for firms.

Figure 17:



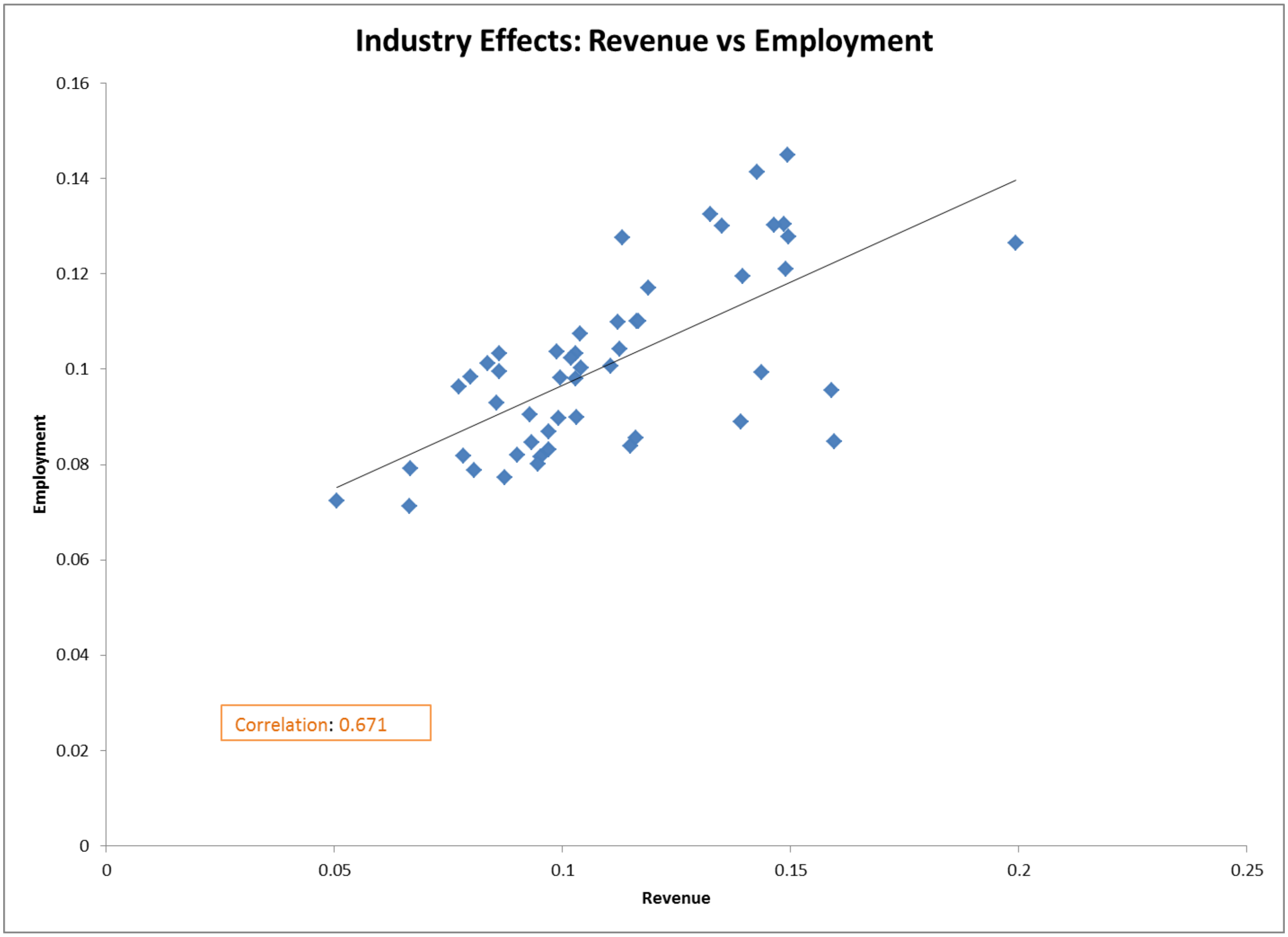
Source: Statistics computed from the Revenue enhanced LBD subset 1998-2000, 2003-2011. Reported are estimated effects of linear probability models on industry effects.

Figure 18:



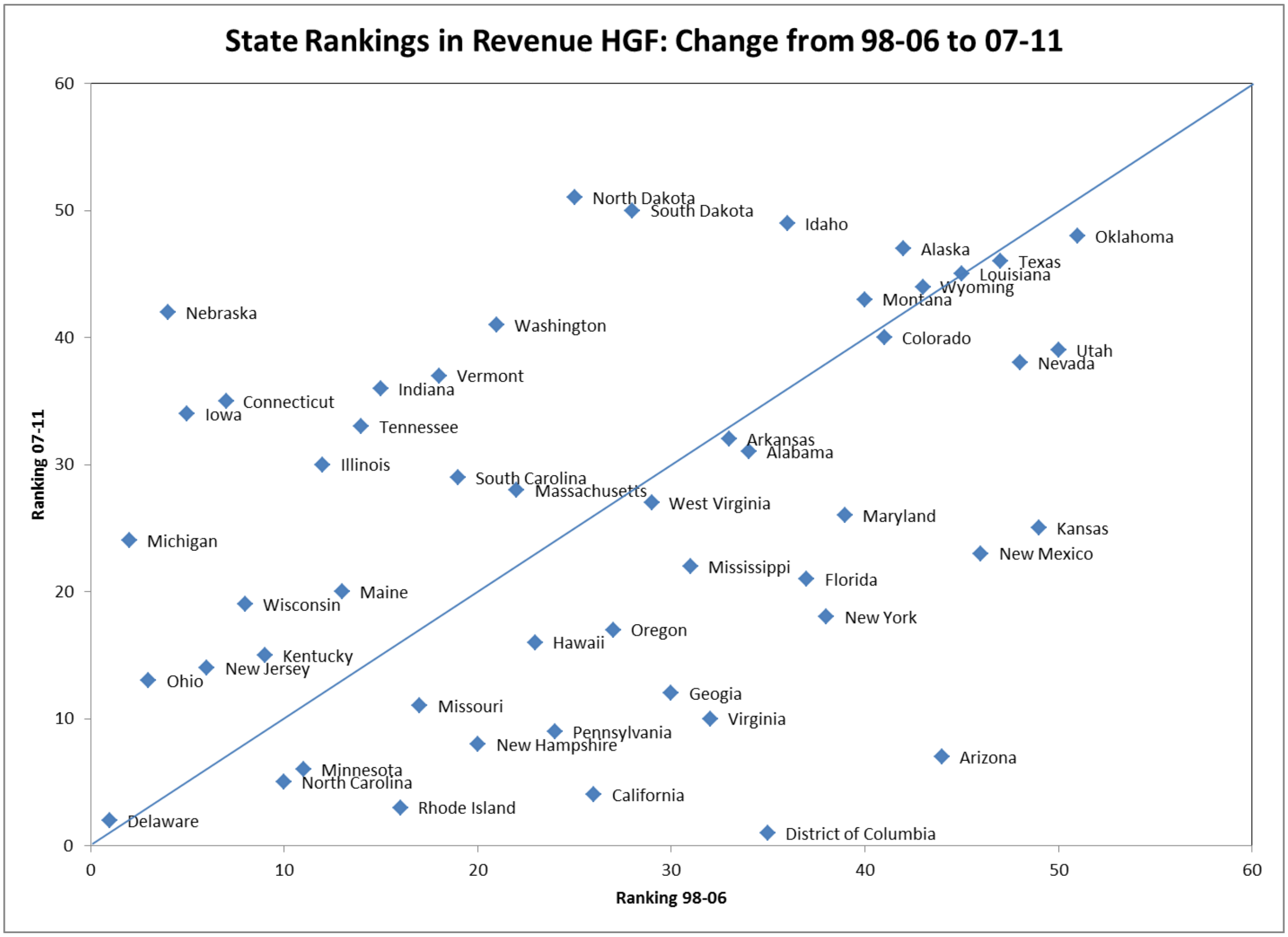
Source: Statistics computed from the Revenue enhanced LBD subset 1998-2000, 2003-2011. Reported are estimated effects of linear probability models on industry effects.

Figure 19:



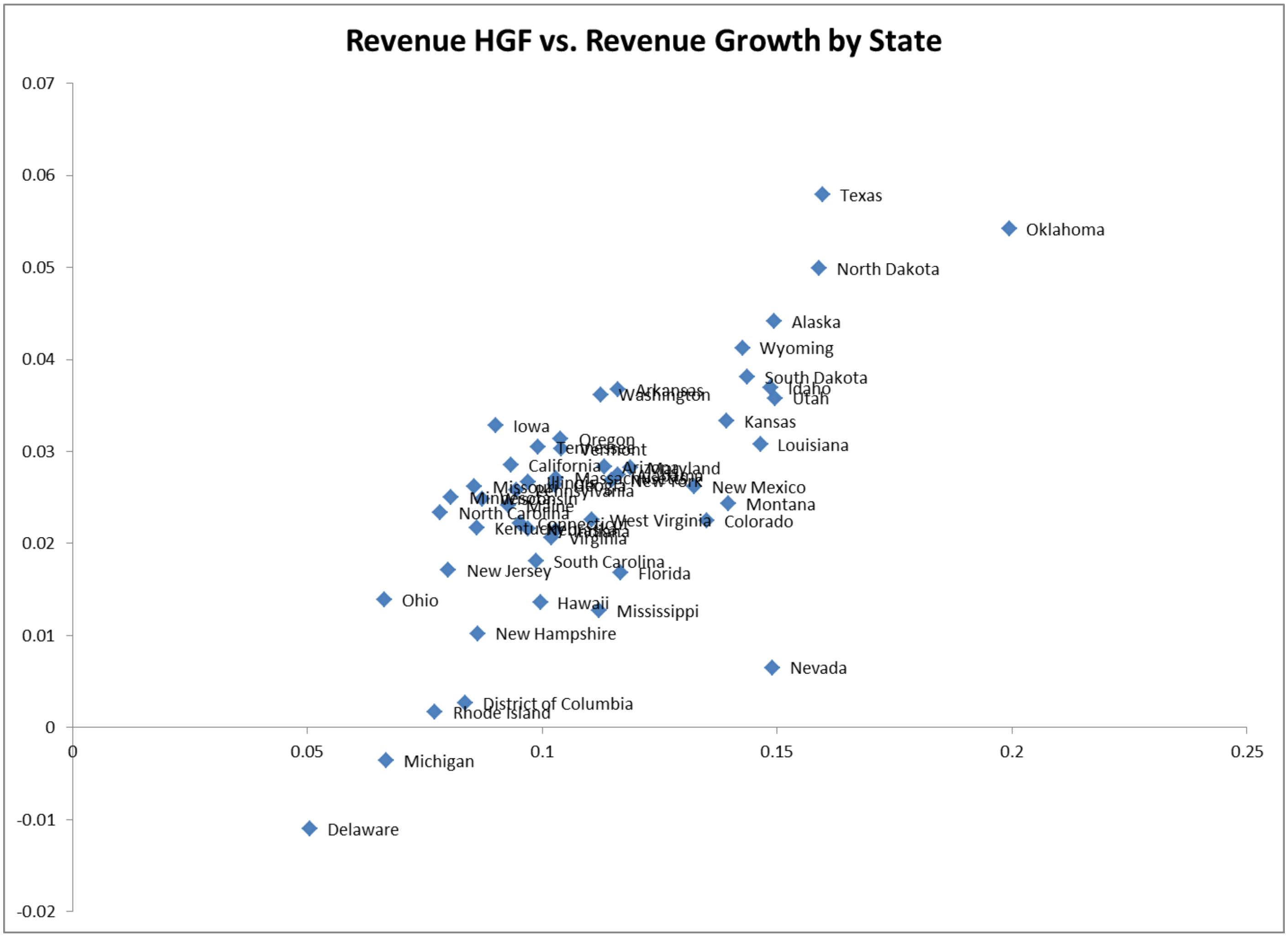
Source: Statistics computed from the Revenue enhanced LBD subset 1998-2000, 2003-2011. Reported are estimated effects of linear probability models on state effects.

Figure 20:



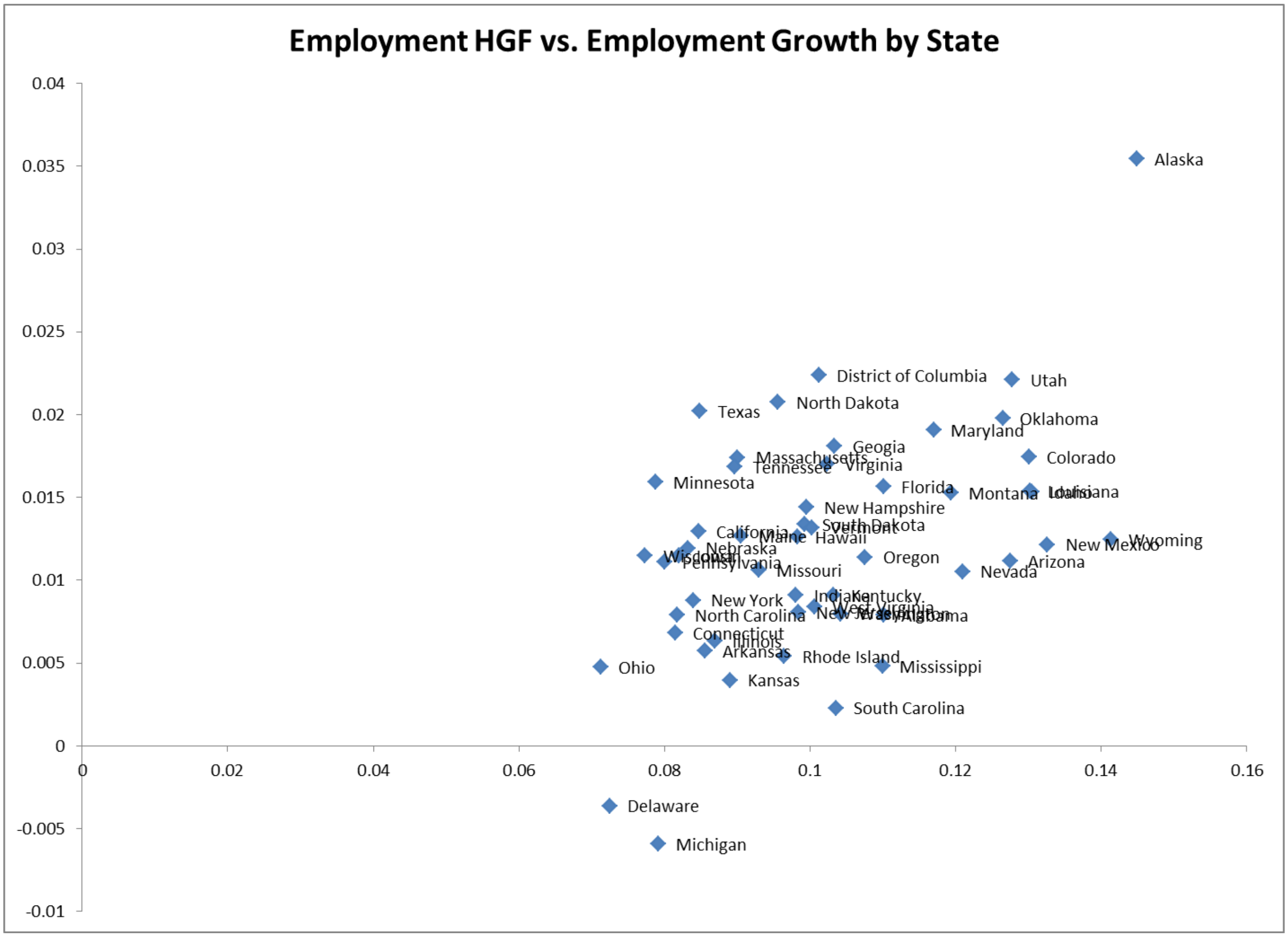
Source: Statistics computed from the Revenue enhanced LBD subset 1998-2000 and 2003-2011. The rankings for 1998-06 use the estimates from the 1998-06 period (except for 2001 and 2002) and the rankings of 2007-11 use the estimates from the 2007-11 period. Reported are estimated effects of linear probability models on state effects.

Figure 21a:



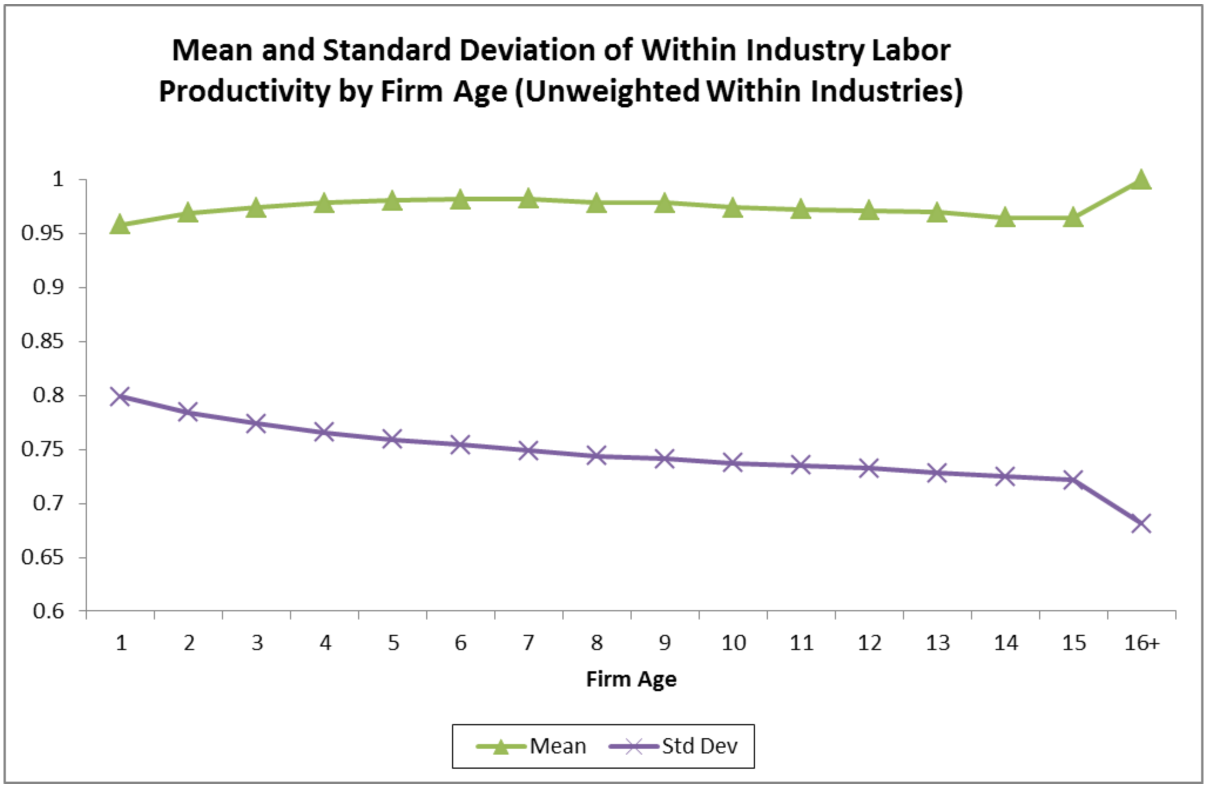
Source: Statistics computed from the Revenue enhanced LBD subset 1998-2000, 2003-2011. Reported Revenue HGF are estimated effects of linear probability models on state effects. Mean Revenue Growth is revenue weighted mean revenue growth for firms.

Figure 21b:



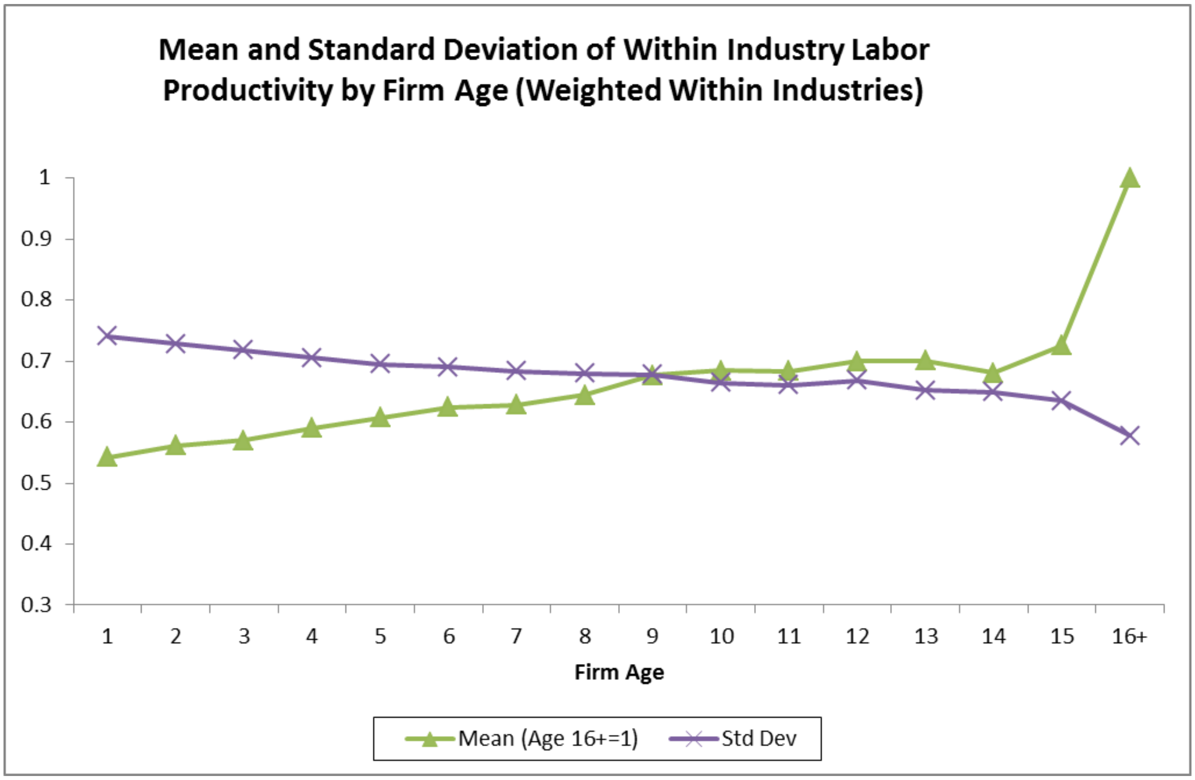
Source: Statistics computed from the Revenue enhanced LBD subset 1998-2000, 2003-2011. Reported Employment HGF are estimated effects of linear probability models on state effects. Mean Employment growth is employment weighted mean employment growth for firms.

Figure 22a:



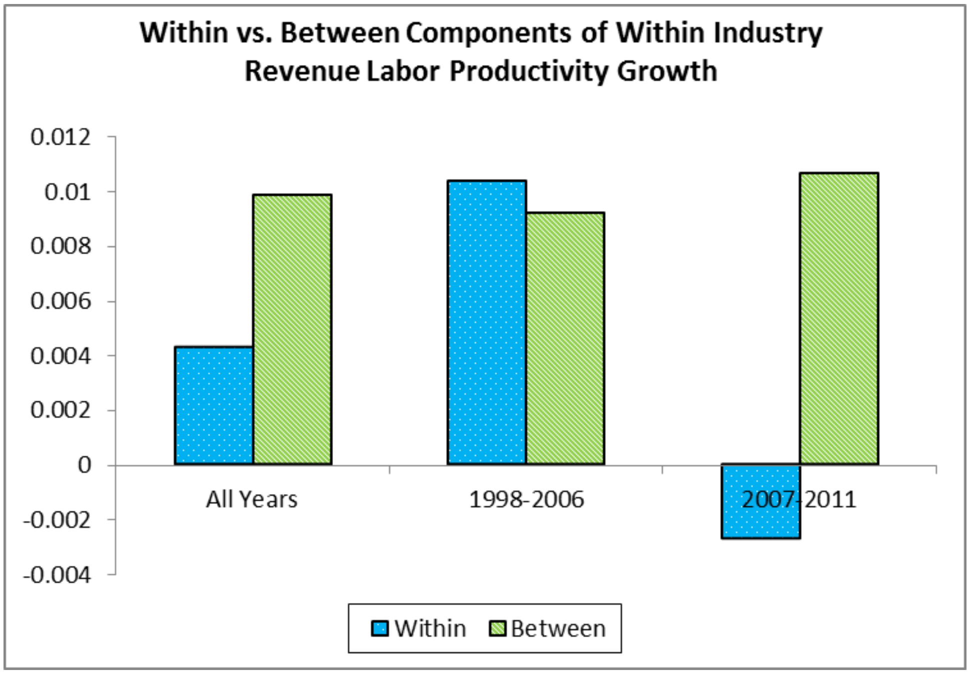
Source: Statistics computed from the Revenue enhanced LBD subset 1998-2000 and 2003-2011.

Figure 22b:



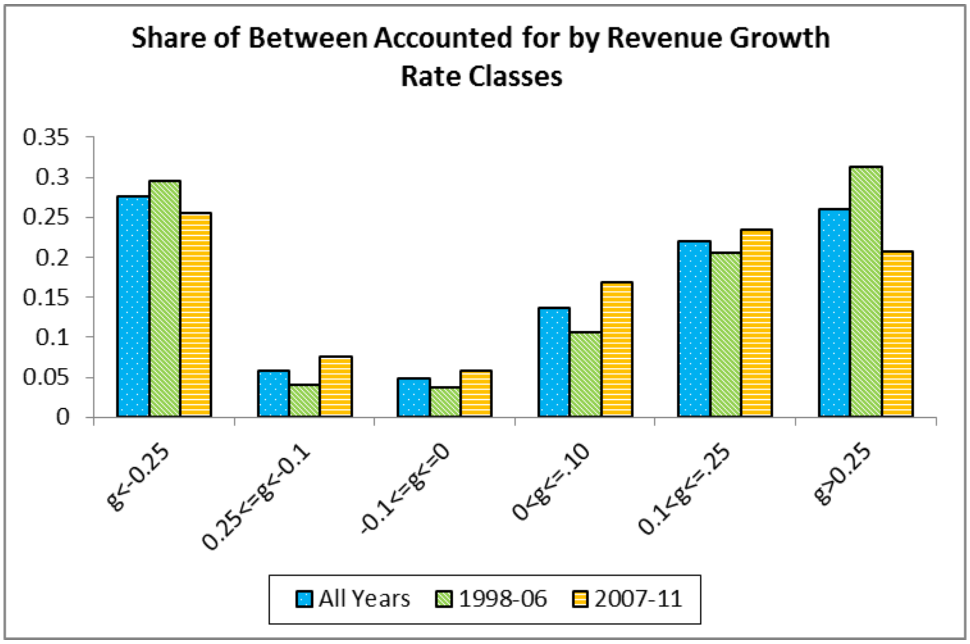
Source: Statistics computed from the Revenue enhanced LBD subset 1998-2000 and 2003-2011.

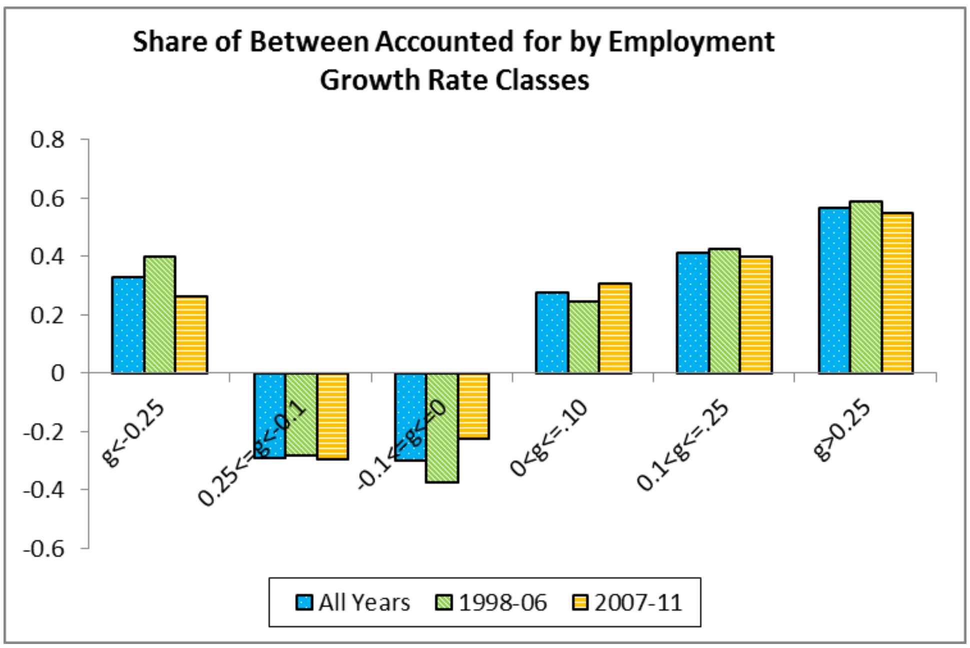
Figure 23



Source: Statistics computed from the Revenue enhanced LBD subset 1998-2000 and 2003-2011.

Figure 24:





Source: Statistics computed from the Revenue enhanced LBD subset 1998-2000 and 2003-2011.