Innovation, Productivity Dispersion, and Productivity Growth

Lucia Foster, Cheryl Grim, John Haltiwanger, and Zoltan Wolf[1]

April 7, 2018

Abstract

We examine whether underlying industry innovation dynamics are an important driver of the large dispersion in productivity across firms within narrowly defined sectors. Our hypothesis is that periods of rapid innovation are accompanied by high rates of entry, significant experimentation and, in turn, a high degree of productivity dispersion. Following this experimentation phase, successful innovators and adopters grow while unsuccessful innovators contract and exit yielding productivity growth. We examine the dynamic relationship between entry, productivity dispersion, and productivity growth using a new comprehensive firm-level dataset for the U.S. We find a surge of entry within an industry yields with a lag an increase in productivity *dispersion* and then after a subsequent lag an increase in productivity *growth*. These patterns are more pronounced for the High Tech sector where we expect there to be more innovative activities. These patterns change over time suggesting other forces are at work during the post-2000 slowdown in aggregate productivity.

[1] Foster and Grim: Center for Economic Studies, U.S. Census Bureau; Haltiwanger: University of Maryland; Wolf: Westat. John Haltiwanger was also a part-time Schedule A employee of the Census Bureau at the time of the writing of this paper. This is a revised version of a paper presented at the NBER/CRIW conference on *Measuring and Accounting for Innovation* in the 21st Century held in Washington, D.C. in March 2017. We thank Nick Bloom, Jonathan Haskel, Ron Jarmin, Jim Spletzer, and CRIW Conference attendees for helpful comments on an earlier draft. Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed.

1. Introduction

The large within-industry productivity *dispersion* commonly found in the firm-level productivity literature (Syverson (2011)) may reflect many factors and mechanisms: idiosyncratic productivity shocks, managerial ability and practices, frictions, distortions, degree of competition, economies of scope, and product differentiation. In healthy economies, reallocation of resources away from low productivity to high productivity firms acts to reduce this dispersion and yields productivity growth. Thus, it is already well understood that within-industry productivity dispersion and productivity growth are related. In this paper, we explore a related but distinct hypothesis relating within-industry productivity dispersion and productivity growth in the context of innovation dynamics within industries.

We investigate this hypothesis in the context of the surge in U.S. productivity in the 1990s to early 2000s and the subsequent productivity slowdown (Fernald (2014), Byrne, Sichel and Reinsdorf (2016), and Andrews et al. (2016)). Some have hypothesized that this reflects a slowdown in the pace and implementation of innovation and technological change especially in the IT intensive sectors (Gordon (2016) and Byrne, Oliner and Sichel (2013)). Others have argued that there is an increase in frictions and distortions slowing down productivity enhancing reallocation dynamics (e.g., Decker et al. (2018)) or the diffusion in productivity (Andrews et al. (2016)).

Our focus is not on the productivity surge and slowdown per se but rather to take a step back to investigate the dynamics we observe between entry, productivity dispersion, and productivity growth using firm-level data. For this purpose, we use a new economy-wide dataset tracking entry, productivity dispersion, and productivity growth at the firm-level. We are especially interested in the hypothesized role of innovation and technological change in these dynamics.

We pursue an indirect approach rather than a direct approach to measuring innovation. Directly measuring innovation is a challenge; much of the focus in the literature focuses on measuring *inputs* to innovation (such as R&D expenditures) or proxies for the *output* of innovation (such as patents). While much progress has been made using these more direct approaches, it is likely that such measures capture only a small fraction of firm innovative activity. In our approach, we attempt to indirectly measure the presence of innovation through its impact on more easily measured concepts (entry and productivity). This is analogous to the approach in astronomy of measuring black holes through the characteristics of nearby visible stars whose properties act according to laws of nature. While social science does not have laws of nature as such, economics has organizing principles about the behavior of economic agents that help direct our attention to areas of the economy where innovation is likely to have occurred. Our objective is to explore this indirect approach with some novel empirical analysis and in turn to discuss questions that can be addressed with these and related data.

A useful starting point for our analysis is Gort and Klepper (1982) who hypothesize stages of firm dynamics in response to technological innovations. While they focus on product innovations, their insights apply to process innovations as well. An insight key for our purposes is that periods of rapid innovation yield a surge in entry, a period of significant experimentation, followed by a shakeout period when successful developers and implementers grow while unsuccessful firms contract and exit. This is relevant for productivity dispersion because the success or failure of entrepreneurs in the process of experimentation can contribute to dispersion, the subsequent reallocation of resources, and eventually, economic growth.

A large subsequent literature has developed models of innovation via creative destruction with some of these features (see, e.g., Jovanovic (1982), Klette and Kortum (2004) and Lentz and Mortensen (2008)). Related theoretical models that highlight the role of entrants and young firms for innovation in models of creative destruction include Acemoglu et al. (2017). These creative destruction models of innovation are related to the empirical literature that finds the reallocation of resources is an important determinant of aggregate productivity growth (Griliches and Regev (1992); Baily, Hulten, and Campbell (1992); Baily, Bartelsman, and Haltiwanger (2001); Petrin, White, and Reiter (2011); Foster et. al. (2017)). Also related to these ideas are the now well-known findings that young businesses, particularly those in rapidly growing sectors, exhibit substantial dispersion and skewness in the growth rate distribution (Dunne, Roberts and Samuelson (1989); Davis, Haltiwanger and Schuh (1996); Haltiwanger, Jarmin and Miranda (2013); Decker, Haltiwanger, Jarmin and Miranda (2016)).

An underexplored empirical area of research is the evolution of the productivity distribution within the context of these innovation dynamics. Partly this has been due to data limitations. For example, Gort and Klepper (1982) investigated their hypotheses mostly on firm-level registers that permitted tracking entry, exit, and continuers in industries but not outcomes like productivity growth and dispersion. While there has been an explosion of research using firm-level data since then, much of what we know about productivity dispersion and dynamics is about the manufacturing sector (Syverson (2011)). We overcome these data limitations by exploiting a newly developed *economy-wide* firm-level database on productivity (Haltiwanger, Jarmin, Kulick and Miranda (2017)). Using this database, we investigate these issues focusing on the nature of the relationship between industry productivity growth and within-industry productivity dispersion. We also look at the relationship between firm dynamics and the evolution of the firm-level productivity dispersion in industries undergoing rapid productivity growth.

To preview our results, we first report broad patterns in aggregate and micro data that provide additional motivation for our analysis. We show that the period prior to 2000 has rising entry, increased within-industry productivity dispersion, and high productivity growth in the High Tech sectors of the U.S. economy. In contrast, the period following 2000 has falling entry, increased within-industry productivity dispersion, and low productivity growth in the High Tech sectors. We also find that within-industry dispersion in productivity is much greater for young compared to mature firms. These findings are not novel to this paper (see Decker et al. (2016, 2018)) but serve a useful backdrop for our analysis.

To help understand these broad-based patterns, we use firm-level data for the entire U.S. private sector to construct measures of firm entry, within-industry productivity dispersion, and industry level productivity growth at a detailed industry level. We use low frequency variation to abstract from high frequency cyclical dynamics and a difference-in-difference specification that controls for time and industry effects. Using this specification, we find that a surge in entry in an industry is followed by a rise in within-industry productivity dispersion and a short-lived slowdown in industry-level productivity growth. Following this, there is a decline in productivity dispersion but an increase in productivity growth. These findings are larger quantitatively for industries in the High Tech sectors of the U.S. economy.

To better understand the role innovation plays in the reallocation of resources between firms, we also use these data to explore the contribution of reallocation dynamics to productivity growth. We find that underlying the productivity surge in the High Tech sectors in the late 1990s is a high contribution of increased within-industry covariance between market share and productivity. The productivity slowdown in the post-2000 period in High Tech is due to both a decrease in within-firm productivity growth but also a decrease in this covariance.

These findings are broadly consistent with the Gort and Klepper (1982) hypotheses that a period of innovation yields a period of entry and experimentation followed by a shakeout period with successful firms growing and unsuccessful firms contracting and exiting. In this respect, some aspects of our results provide confirming micro-level evidence for the hypothesis that the productivity slowdown is due to a decreased pace of innovation and technological change. However, we are reluctant to make that inference for at least two reasons. First, our investigation does not include direct measures of innovation. Second, the patterns in the post-2000 period are not consistent with a slowdown in innovation as the primary source for the post-2000 productivity slowdown. With that hypothesis, we would have expected to observe a decline in productivity dispersion; instead, the findings in Decker et al. (2018) show that dispersion is rising even though the fraction of activity accounted for by young firms is falling dramatically in the post-2000 period.[2]

We view the results from our empirical exercises as suggestive highlighting the potential measurement benefits of studying the joint dynamics of entry, productivity dispersion, and productivity growth. In the second half of the paper, we discuss open questions and next steps suggested by our analysis with a focus on the measurement and analysis of innovation.

The rest of the paper proceeds as follows. In the next section, we provide more discussion on the conceptual underpinnings for our empirical analyses and interpretations. We describe the data and measurement issues in Section 3. Our empirical exercises examining patterns of entry, productivity dispersion, and productivity growth and reallocation dynamics are in Section 4. In Section 5, we discuss open questions, measurement challenges and areas for future research suggested by our analysis. Section 6 presents concluding remarks.

2. Conceptual Underpinnings

We begin by reviewing the sources of measured productivity dispersion within industries. For this purpose, it is critical to distinguish between underlying sources of technical efficiency and measured productivity across firms in the same sector. In empirical applications, the latter is typically some measure of "revenue productivity," which sometimes is a multi-factor measure of

[2] There are additional reasons to be cautious in this inference. Decker et al. (2018) find that there has been a decrease in responsiveness of growth and exit to productivity growth. The latter is consistent with an increase in adjustment frictions. We discuss these issues further below.

input and other times is revenue per unit of labor. In either case, revenue productivity measures are inherently endogenous to many different mechanisms and factors. For ease of discussion, we follow the recent literature in referring to measures of technical efficiency as TFPQ, revenue measures of total factor productivity as TFPR, and revenue-based measures of labor productivity as LPR.

As mentioned in the Introduction, measured productivity differences between firms may be attributed to a variety of factors. The link that connects innovative activity and dispersion is experimentation following an innovation that generates heterogeneity in the factors that cause dispersion. Many models of firm heterogeneity start with the premise that there is some source of exogenous as well as endogenous differences in TFPQ across firms. In some models this is due to inherent characteristics of the firm reflecting permanent differences in the technology distribution (e.g., Lucas (1978) and Jovanovic (1982)) that may in turn stem from many factors such as managerial ability. In other models, the firms are subject to new, and typically persistent, draws of TFPQ each period (Hopenhayn (1992), Hopenhayn and Rogerson (1993) and Ericson and Pakes (1995)). Endogenous differences in TFPQ may stem from differences in adoption of management practices (e.g., Bloom et. al. (2017)) or differences in realizations of endogenous innovations in investment (e.g., Acemoglu et. al. (2017)).[3] A variety of reasons have been put forth to justify how high and low TFPQ firms can coexist (i.e., why the most productive firms do not take over the market); these range from economies of scope (Lucas (1978)) to product differentiation (Melitz (2003)) to adjustment frictions (Hopenhayn and Rogerson (1993) and Cooper and Haltiwanger (2006)) and all of these factors likely play some role empirically.

These factors, together with the ample evidence that there is price heterogeneity within sectors (Syverson (2004), Foster, Haltiwanger, and Syverson (2008), and Hottman, Redding and Weinstein (2016)), imply that revenue productivity (TFPR and LPR) dispersion will also be present within sectors and revenue productivity measures will be correlated with TFPQ at the

[3] Even in these cases with endogenous adoption of management practices or investment in innovation there still is typically an underlying exogenous source of heterogeneity that induces heterogeneous adoption/investment and/or there are stochastic returns from such adoption/investment. These potentially endogenous factors play an important role. For example, using data from the Management and Organizational Practices Survey, Bloom et al. (2017) find large differences in management practices across plants and that these differences explain about one-fifth of the difference between the 90th and 10th productivity percentiles.

firm level (Haltiwanger (2016) and Haltiwanger, Kulick and Syverson (2018)).[4] Thus, one source of variation in measured revenue productivity across sectors and time is variation in dispersion in TFPQ as well as other idiosyncratic shocks to fundamentals such as demand shocks. Another factor that impacts within-industry revenue productivity dispersion is the business climate as broadly defined to include distortions in output and input markets that impede growth at more productive firms and contraction and exit at less productive firms. This has been the theme of the recent misallocation literature (Restuccia and Rogerson (2008), Hsieh and Klenow (2009), and Bartelsman et al. (2013)). An economy or industry that experiences a deterioration in the business climate, should from this perspective, exhibit a decline in productivity along with a rise in dispersion in revenue productivity. The intuition is that rising frictions and distortions reduce the tendency for marginal revenue products to be equalized implying in turn a rise in revenue productivity dispersion. A detailed discussion on how these factors affect dispersion in revenue-based productivity measures can be found in Foster et al. (2016).

As another example, Hurst and Pugsley (2011, 2017) emphasize that non-pecuniary benefits play an important role in the occupational decision to become an entrepreneur. Their insight is that the potentially different incentives underlying entrepreneurial behavior will be reflected by measured differences in productivity as well as firm size and growth. They argue that there are large differences across sectors in terms of attractiveness for entrepreneurs with high non-pecuniary benefits. This might be yet another factor that helps account for dispersion in TFPQ within some industries but this is likely less important in innovative intensive industries. Such sectoral heterogeneity is one of the (many) reasons we control for detailed industry fixed effects in our empirical analysis.

How do innovation and firm dynamics associated with innovation relate to heterogeneity in measured productivity? The basic idea in Gort and Klepper (1982) is that a period of intensive transformative innovation within an industry is accompanied by (and/or induces) entry. This period is characterized by entrants engaging in substantial experimentation and learning. Since

[4] There is a knife-edge case emphasized by Hsieh and Klenow (2009): with constant returns to scale and isoelastic demand without adjustment costs or other factors (like overhead labor), revenue productivity should have zero dispersion in equilibrium even if there is dispersion in TFPQ. The reason is the elasticity of firm level prices with respect to TFPQ is equal to exactly -1 in this knife-edge case (see Haltiwanger, Kulick and Syverson (2018)). This knife-edge case is interesting theoretically to help fix ideas, but is not very useful empirically since there is much evidence that factors such as adjustment costs make this knife-edge case irrelevant in practice.

experimentation and learning involves trials and errors which yield different outcomes, there is likely to be an increase in dispersion in TFPQ accompanied by increases in dispersion in TFPR and LPR. In addition, if increased innovation and entry beget a higher share of young businesses and they are more likely to face more frictions, uncertainty, and distortions, then dispersion in TFPR and LPR will rise further.[5] As the experimentation phase identifies successful innovators and adopters of new products and processes, these firms are likely to grow; while their unsuccessful competitors will contract and exit. This process leads to period of productivity growth from both within-firm productivity growth (at successful innovators) as well as productivity enhancing reallocation dynamics. The latter should reduce productivity dispersion both through selection but also the maturing of the more successful firms.

With the above considerations in mind, we hypothesize that the innovation dynamics described in Gort and Klepper (1982) imply the following about entry, productivity dispersion and productivity growth dynamics. Following a surge in entry accompanying innovation, we should observe a period of rising dispersion in LPR within industries that will in turn be followed by increased industry-level productivity growth. The latter will reflect both within-firm productivity growth of the successful developers and adopters and the reallocation of resources to such firms.

The Gort and Klepper (1982) hypothesized impact of innovation dynamics on productivity dispersion is not inconsistent with the many different sources of productivity dispersion discussed above. Instead, we view the Gort and Klepper hypothesis as providing insights into potentially important driving forces for low frequency within industry variation in the dynamic relationship between innovation, entry, productivity dispersion and productivity growth. The empirical analysis that follows focuses on these low frequency within industry dynamics.

3. Data and Measurement

[5] We use the term “frictions” for factors that the social planner cannot overcome such as adjustment costs that are part of the technology of adjustment. In contrast, we use the term “”distortions” for market failures, policies, or institutions that impede firms adjusting to their optimal size. Jovanovic (1982) provides motivation for greater information frictions faced by young firms. Hsieh-Klenow (2009) provide motivation for why young firms are likely to face higher distortions due, for example, to imperfect capital markets.

Our main dataset in this paper is a newly developed extension to the Longitudinal Business Database (LBD). The LBD is an economy-wide establishment-level database primarily derived from the Census Bureau's Business Register and augmented with other survey and administrative data (see Jarmin and Miranda (2002)). It covers the universe of employer businesses in the non-farm business sector of the U.S.: about 7 million establishments and 6 million firm observations per year for 1976-2013. It contains establishment-level information on detailed industry, geography, employment, and parent firm affiliation. The LBD has robust links for businesses over time making this dataset particularly well-suited for the measurement of business dynamics such as job creation and destruction, establishment entry and exit, and firm startups and shutdowns. These links make it possible to aggregate the establishment-level data to the firm-level where firm growth dynamics abstract from mergers and acquisitions and other ownership activity.

A firm startup is defined as a new firm entity with all new establishments; a firm exit is defined as a firm entity that ceases to exist with all of its establishments shutting down; and firm growth is measured as the employment weighted average of the establishments owned by the firm (for details see Haltiwanger, Jarmin and Miranda (2013)). These features also make it feasible to define firm age in a manner that abstracts from mergers and acquisitions and ownership change activity. A firm's age is determined by its longest-lived establishment at the time of the firm's founding and then progresses one additional year over calendar time. Firm-level industry is measured as the modal industry for the firm based on its employment shares across 6-digit or 4-digit NAICS industries. In this analysis, we focus on 4-digit NAICS industries.[6]

We do not use direct measures of innovation in our empirical analysis; instead we use a surge of entry and young firm activity as an indirect proxy for innovative activity. Gort and Klepper (1982) suggest that Stage 1 of a period of increased within-industry transformative innovation is followed by a surge of entry (Stage 2). To shed further light on this process, we group industries into High Tech and other industries (which we call Non Tech). For High Tech,

[6] One concern is this definition of industry is a potential source of measurement error for large, complex multi-units especially since much of our analysis exploits within industry variation in productivity dispersion and growth. The use of 4-digit as opposed to 6-digit industry effects mitigates this concern somewhat. Decker et al. (2018) have explored this issue using a more sophisticated approach to controlling for industry-year effects (based on taking into account the full distribution of employment shares for each firm) and found that the patterns of dispersion and growth within industries are largely robust to this concern.

we adopt the strategy of Decker et al. (2018) who follow Hecker (2005) in defining High Tech industries as the Science, Technology, Engineering, and Math (STEM) intensive industries. In practice, High Tech industries include all of the standard Information and Communications Technology (ICT) industries as well as biotech industries.

Our productivity measure relies on the recently developed firm-level measures of nominal revenue in the LBD developed by Haltiwanger, Jarmin, Kulick, and Miranda (2017) (hereafter HJKM). HJKM use nominal revenue data at the tax reporting or employer identification number (EIN) level from the Business Register (the underlying source for the LBD) to create measures of nominal revenue for over 80 percent of firms in the LBD for their sample period. To mitigate issues of selection due to missingness, they develop inverse propensity score weights so that the revenue sample is representative of the full LBD. We use the HJKM revenue enhanced LBD in our analysis including the propensity score weights. Following Decker et al. (2018) we convert nominal revenue to real measures using BEA price deflators at the industry level (this involves using 4-digit deflators when available and 3 or even 2-digit deflators otherwise). Our firm-level measure of labor productivity (hereafter, "productivity") is the log of the ratio of real revenue to employment. A key limitation of this measure is that the output concept is a gross concept rather than value-added so is not readily comparable across industries (see HJKM). Following HJKM and Decker et al. (2018), we focus on patterns controlling for detailed (4-digit) industry and year effects.[7] We provide further details about this in our empirical exercises below.

Our econometric analyses are based on industry/year-specific moments of firm-level productivity. We construct within-industry measures of productivity dispersion and within-industry measures of productivity growth and supplement these with industry-level information on start-up rates from the full LBD. We tabulate measures such as the share of employment accounted for by young firms (young is less than 5 years old) and the share of employment accounted for by startups (firm age equal to zero). The version of the LBD we use covers 1976-2013 so we can construct these measures for years prior to the available revenue data (available

[7] HJKM and Decker et al. (2018) use 6-digit NAICS as compared to our use of 4-digit NAICS. We use the latter for two reasons. First, this mitigates the measurement problems of using modal industry. Second, the focus of our analysis is industry-level regressions using moments computed from the firm-level data. The 6-digit NAICS data are quite noisy for industry-level analysis particularly analysis that is not activity weighted.

1996-2013). This facilitates some of the dynamic specifications that use lagged entry rates in our analysis below.

Our dispersion measure throughout this paper is the interquartile range (IQR) within an industry in a given year. We focus on the IQR because it is less sensitive to outliers than the standard deviation (see Cunningham et al. (2017)). Our measure of within-industry productivity growth aggregates real revenue and employment data to the 4-digit industry level and then we compute the log first difference at the industry-level. In our exercises using the Dynamic Olley-Pakes decomposition developed by Melitz and Polanec (2015), we exploit firm level changes in labor productivity as well as the other terms in that decomposition.

Finally, the focus of this paper is on the longer-term relationship between these three important concepts of entry, productivity dispersion, and productivity growth. We have two strategies to attempt to abstract away from business cycle variation. In some exercises we use Hodrick-Prescott (HP) filtering to ameliorate the impact of business cycles; in other exercises we use 3-year non-overlapping periods to conduct our analysis.

4. Empirical Evidence

We examine the relationship between innovation, entry, productivity dispersion, and productivity growth motivated by the hypotheses in Gort and Klepper (1982) (GK hereafter) discussed in Section 2. Assuming that the businesses in High Tech industries are more innovative than in other sectors, these hypothesized GK dynamics should be more likely to occur in these industries. We explore whether this is the case in our data by examining whether the nature of the dynamics differs between High Tech and Non Tech industries. We start our empirical analysis by providing some basic facts about the patterns of entry, productivity dispersion, and productivity growth for industries grouped into the High Tech and Non Tech sectors. These basic facts are already reasonably well-known in the literature but they provide helpful motivating evidence for our subsequent analysis.

4.1 Basic facts

We start with the key industry-level indicator concerning startups and the share of activity accounted for by young firms. In Figure 1 we plot the employment shares for High Tech (green) and Non Tech (red) industries for both startups (Panel A) and young firms (Panel B). There are noticeable differences in the startup patterns for High Tech as compared to Non Tech in Panel A. While Non Tech shows a gradual decline over time in employment shares, High Tech shows a humped shape pattern culminating in the three-year period between 1999 and 2001. This difference is even more dramatic for young firms as is shown in Panel B of Figure 1. Together these panels highlight the surge in entry and young firm activity in High Tech in the 1990s.[8]

We next turn to the second key moment of interest: within-industry productivity dispersion. We start by simply examining the within-industry dispersion of productivity for firms based on their age (Young versus Mature) and whether they are in High Tech or Non Tech. Again, dispersion is measured by the interquartile range within an industry in a specific year. We use the same time-invariant industry employment weights to aggregate the industry level patterns to High Tech and Non Tech industries. Figure 2 plots dispersion for Young (solid lines) and Mature (dashed lines) and High Tech (green) and Non Tech (red).[9]

As expected, Young firms (regardless of their Tech status) have more dispersion within industries than Mature firms (solid lines are well above the dashed lines). The differences between High Tech and Non Tech vary over the two firm-age groups. For Young Firms, High Tech generally has greater dispersion that Non Tech; for Mature Firms, Non Tech has greater dispersion than High Tech. This difference is a reminder that there are many things driving dispersion and Non Tech is a heterogeneous group. Moreover, within firm age groups, dispersion is rising throughout the whole sample period. This is consistent with GK hypotheses of more experimentation; however, it is also consistent with potentially rising frictions for young firms leading to greater dispersion in productivity.[10]

Finally, we examine labor productivity growth at the aggregate (broad sector) level from official Bureau of Labor Statistics (BLS) statistics and aggregates using our micro-level data. We start by focusing on BLS data. Panel A of Figure 3 plots BLS labor productivity growth rates for

[8] The patterns in Figure 1 are already well known (see, e.g., Haltiwanger, Hathaway and Miranda (2014) and Decker et al. (2016)).
[9] Note that this figure is similar to analysis conducted in Decker et al. (2018), see their Figure 7. The latter controls for 6-digit industry effects. Also, Decker et al. (2018) use a more sophisticated manner of controlling for such effects for multi-unit establishment firms that have activity in more than one 6-digit industry. The patterns we show in Figure 2 are consistent with these alternatives suggesting our use of 4-digit industry effects is not distorting the patterns.
[10] Decker et al. (2018) explore the hypothesis whether rising dispersion is due to rising frictions/distortions and focus on declining responsiveness to shocks as one potential explanation. We return to discussing this issue further below.

the High Tech and Non Tech broad sectors measured as employment-weighted within-industry (4-digit) labor productivity growth based on gross output per worker. For employment weights, we use time-invariant employment shares so the depicted patterns hold industry composition constant. We present four measures in this panel: the annual BLS labor productivity growth (dashed lines) and the smoothed HP filtered version of this growth (solid lines) for High Tech (green) and Non Tech Industries (red). It is evident from the annual versions of the plots (dashed lines) that there is substantial cyclicality. Turning to the HP filtered (solid) lines, we see rising productivity growth in High Tech and then falling sharply post 2000 confirming earlier studies. Non Tech has much more muted patterns but slight rise in the 1990s and falling in post 2000.

Next we look at the aggregates constructed from the firm-level data and compare them to the BLS data. The micro aggregates are based on employment-weighted within-industry labor productivity growth measured by log real gross output per worker. That is, using the firm-level data we first construct industry-level labor productivity growth and then use the same type of time- invariant industry employment weights to aggregate to High Tech and Non Tech sectors. Panel B of Figure 3 plots the HP filtered labor productivity growth rates for BLS aggregate data (solid lines, repeating those from Panel A) and Census micro data (dashed lines) for High Tech (green) and Non Tech (red). We find that micro based aggregates track BLS productivity patterns reasonably well.

4.2 Dynamic Relationship between Entry and Productivity

To explore the dynamic relationship between entry, productivity dispersion, and productivity growth, we use a panel regression specification exploiting industry-level data (from micro-level data) over time using a standard difference-in-difference approach. The hypotheses from GK are that a surge of within industry entry will yield an increase in dispersion followed by an increase in productivity. To investigate these hypotheses, we estimate the following specification:

$$Y_{is} = \lambda_s + \lambda_i + \sum_{k=1}^{2}[\beta_k \, Tech * Entry_{is-k} + \delta_k \, NonTech * Entry_{is-k}] + \varepsilon_{is}, \quad (1)$$

where Y_{is} denotes either within-industry/year productivity dispersion or within-industry productivity growth. Since we are primarily interested in low frequency variation, we calculate productivity growth as the three-year average for subperiods in our sample (1997-99,…,2009-2011,2012-13; note that the last period is only two years). We use a standard difference-in-difference specification with period effects (λ_s) and industry effects (λ_i). The *Tech* dummy is equal to one if industry is in High Tech and is 0 otherwise; the *Non Tech* dummy is equal to one if industry is Non Tech and is 0 otherwise. *Entry* is the startup rates from the full LBD. In order to examine the role of lags, we take advantage of the fact that we can measure startups for earlier periods; so we compute startups for the additional three year periods: 1991-1993, 1994-96. We let the impact of entry have a distributed lag form over two three-year subperiods encompassing a total of six years. We view this analysis as exploratory and it would be of interest to consider even richer dynamic specifications that potentially allow for the type of long and variable lags that GK suggest are potentially important.

The results for the specification on productivity dispersion are shown in Table 1. We find that an increase in entry in one sub-period (three-year period) leads to a significant increase in productivity dispersion in the next sub-period. Moreover, this effect is larger in the High Tech sector. The fact that the coefficients on the second lag are not significant suggests that this effect at least diminishes over time. We interpret this to mean that following an innovation (as proxied by the entry rate), there is an increase in productivity dispersion shortly thereafter representing the experimentation and differential success in the development and adoption of innovations.

The analogous results from the productivity growth specification are shown in Table 2. Here there is a different pattern in the timing. An increased in the startup rate results in a decrease in productivity growth in the next sub-period although this effect is only statistically significant in Non Tech. This suggests there is some evidence that the period of experimentation and dispersion can yield an initial drag on productivity. It is only in the subsequent periods, that we see an increase in productivity growth. The productivity growth impact is larger for firms in High Tech industries as compared to firms in Non Tech industries.

The dynamic responses for both productivity dispersion and productivity growth are depicted in Figure 4. While the patterns are more pronounced for High Tech, they are also present for Non Tech. The finding that the entry-to-dispersion-to-growth dynamics are present for industries outside of High Tech suggests that the GK hypothesized dynamics may be more pervasive across a broad range of industries. Further investigation of these issues and the differences in the patterns across industries is an important area for future research.

Given these results, an interesting and open question is whether these dynamics help account for the aggregate patterns of productivity growth and dispersion. Even though more research is needed, we think that the GK dynamics are not sufficient to understand the patterns in

Figures 1-3 for High Tech particularly in the post-2000 period. In High Tech, we observe a rise in entry (Figure 1), a rise in productivity dispersion (Figure 2), and a rise in productivity (Figure 3) in the 1990s. While the timing is not exactly consistent with Figure 4, these 1990s patterns are broadly consistent with GK hypothesized dynamics. However, in the post-2000 period we observe a decline in entry and productivity but a continued and even sharper rise in within-industry productivity dispersion. From the GK perspective, we should have observed a decline in productivity dispersion.

What factors might account for the rising within productivity dispersion in the post-2000 period? Decker et al. (2018) find that there has been declining responsiveness of firms to shocks. They find that high productivity firms are less likely to grow and low productivity firms are less likely to shrink and exit in the post-2000 period relative to earlier periods. They argue that this is consistent with a rise in frictions and distortions and helps explain the decline in productivity and the pace of reallocation in the post-2000 period. They also note that a rise in frictions is consistent with a rise in dispersion in revenue labor productivity as the increase in frictions will slow the pace at which marginal revenue products are equalized. It may also be the case that the same increase in frictions helps account for the decline in entry in the post-2000 period.

For current purposes, this discussion is a reminder that many factors other than innovation dynamics underlie the joint dynamics of entry, productivity dispersion, and productivity growth. In the next section, we explore some of these issues by examining the nature of the contribution of allocative efficiency to productivity growth in High Tech.

4.3 Dynamic Micro-Macro Productivity Decomposition

In principle, the rise in productivity growth following the experimental phase in the GK framework should be due to both within-firm productivity growth of the successful innovators and the reallocation of resources towards such successful innovators. To investigate these issues we use the Dynamic Olley-Pakes (DOP) decomposition developed by Melitz and Polanec (2015).

Melitz and Polanec extend the Olley-Pakes method to include entry and exit in a manner that allows for careful tracking of within-firm changes. Similar to Olley-Pakes, their decomposition of an index of industry-level productivity growth includes terms for changes in average productivity growth and a covariance term, but they split these components out to distinguish between firms that continuously operate and firms that enter and exit. Their decomposition is shown in Equation 2:

$$\Delta P_{it} = \Delta\bar{p}_{it,C} + \Delta cov_C\big(\theta_{ft}, p_{ft}\big) + \theta_{Nt}(P_{Nt} - P_{Ct}) + \theta_{Xt-1}(P_{Ct-1} - P_{Xt-1}) \tag{2}$$

where Δ indicates year-over-year log difference, P_{it} is the index of industry level productivity in industry *i* in period *t* defined as the weighted average of firm level productivity using firm level employment weights θ_{ft} (the share of employment of firm *f* in total industry employment), $\bar{p}_{it}$ is the unweighted average of (log) firm-level productivity for the firms in industry i, C denotes continuer firms (those with employment in both t-1 and t) so that and $Ct-1$ and Ct denote continuers in periods t-1 and t, respectively, Nt denotes entrants from t-1 to t, $Xt-1$ denotes firms that exit from t-1 to t.. The first term in the expression, $\Delta\bar{p}_{it,C}$, represents average (unweighted) within-firm productivity growth for continuing firms; the second term, $\Delta cov_C(\theta_{ft}, p_{ft})$, represents the change in covariance among continuing firms, the third term captures the contribution of entry, while the fourth term captures the contribution of exit.

Using the weighted average of firm-level productivity as an index of industry-level productivity is common in the literature but must be used with appropriate caution. As Decker et al. (2018) show, applying this approach with TFP yields traditional industry-level TFP measures (industry output per composite input) only under constant returns to scale and perfect competition. In the absence of the latter, variations in this index will exhibit greater volatility than traditional measures of TFP as the curvature in the revenue function will yield lower reallocation than is implicit in using this index. Still, this index for industry level labor productivity tracks traditional measures of output per worker relatively well in practice (see Figure 3). Moreover, Appendix B of Decker et. al. (2018) shows that the OP decomposition theoretically tracks aggregate labor productivity more closely than aggregate TFP measures over the empirically relevant range of adjustment costs. We use the DOP decomposition using labor productivity with appropriate caution about interpretation.

In the DOP framework, the changing covariance terms depend critically on (1) there being dispersion in productivity across firms, (2) the covariance between productivity and employment share being non-zero within industries, and (3) the covariance changing over time. We first calculate the components in equation (2) for each industry in each year and then aggregate the annual components to the High Tech level using time-invariant industry employment weights in order to keep industry composition constant (as we have done in Figures 1-3). Focusing on the contribution of the within-industry dynamics in the High Tech sector in this manner helps understand the role of dispersion for this critical innovative set of industries.

Figure 5 reports the annual DOP decomposition where all components are smoothed by the HP filter. We find declining DOP within and covariance terms indicating smaller contributions by both firm level productivity growth and between-firm reallocation. We find only a modest role for net entry but this should be interpreted with caution since this is the average annual net entry contribution reflecting the contribution of entrants in their first year. The contribution of entry arguably takes time and our evidence from Table 2 suggests that this is the case.[11]

We draw several inferences from our related exercises in this section. First, the late 1990s were a period of rapid productivity growth, intensive entry, high young firm activity, rising productivity dispersion --for young firms in particular-- and a large contribution of reallocation activity. Second, the industry level difference-in-difference regressions imply complex timing: entry yields rising productivity dispersion almost shortly after but impacts productivity growth with a significant lag. Third, during the productivity slowdown, entry declined, contributions from within-firm productivity growth were smaller, and reallocation activity declined. Fourth, productivity dispersion kept rising during this period. This last piece of evidence does not mesh well with GK-type dynamics, suggesting that rising dispersion after 2000s appears to be outside the scope of this model. As we have noted above, one possible explanation for the latter is rising frictions and distortions.

5. Conceptual and Measurement Challenges

Our empirical analyses in Section 4 are intended to be exploratory. Our results suggest that there are systematic patterns in the joint dynamics of entry, within-industry productivity

[11] It might seem surprising that the change in the DOP within is so low and then turns negative. Decker et al. (2017) conduct related analysis of the DOP decomposition for the entire private sector. They emphasize that the weighted within term of decompositions such as the Foster, Haltiwanger and Krizan (2001) (FHK) decomposition is larger than the DOP within term. They note that the DOP within is based on unweighted changes in productivity that is dominated by small firms. For the purposes of the current paper this is not a critical issue. In unreported results we have found that the weighted FHK within is larger than the DOP within and always remains positive. However, it declines in the same fashion as the DOP within for High Tech.

dispersion, and within-industry productivity growth that shed light on innovation dynamics. However, there are many open conceptual and measurement questions about using this indirect approach to capturing innovation especially with respect to direct approaches to measuring innovation. In this section, we describe those open questions in light of ongoing and potential measurement and research efforts to understand innovative activity. We do not attempt to provide a survey of the voluminous literature on measuring innovation. Instead, we highlight recent and current efforts with a focus on U.S. statistical agencies in general and the U.S. Census Bureau in particular. As background, the work from two large related research projects at the Census Bureau lies behind much of the indirect approach taken in this paper. The first of these, the LBD project, seeks to improve measures of firm dynamics. The second of these, the Collaborative Micro Productivity (CMP) project, seeks to prove the usefulness of producing higher moment statistics from micro-level data (using productivity as the pilot statistic, see Cunningham et al. (2017)).

In the remainder of this section we discuss four areas of interest for the measurement of innovation and productivity: direct measures of innovation, linking firm entry and innovation, intangible capital measures, and high frequency versus low frequency analysis. We view this section of the paper as relating ongoing efforts to improve measurement in this area to the indirect approach we have taken in the analysis above.

5.1 Direct Measures of Innovation: R&D Expenditures and Patents

Direct measurement of innovation is a challenge. In this subsection, we highlight a few approaches to this challenge that are particularly relevant and feasible. Some of these activities are part of ongoing research projects underway at Census. One common approach is to measure *inputs* to innovative activity such as R&D expenditures. The Census Bureau conducts the Business Research and Development and Innovation Survey (BRDIS) in accordance with an interagency agreement with the National Science Foundation (NSF). The BRDIS (or its predecessor, the Survey of Industrial Research and Development (SIRD)) has collected firm-level information on R&D expenditures since 1953. Griliches (1980) was one of the first users of this micro data from the SIRD (combining it with other Census datasets). Since then, the survey has expanded in scope from its original focus on large manufacturing companies. However, there remain many challenges with using R&D expenditures to proxy innovation activities (see Foster, Grim, Zolas (2016)) and we focus on a few of the more relevant challenges. First, measurement quality may vary over sectors of the economy because measuring R&D expenditures is more complicated in some sectors (for example, the Service and Retail sector; see, Brown et. al (2005)). Second, measurement quality may vary over types of firms because activities that constitute R&D and innovation activity are easier to capture in large, mature firms with dedicated R&D divisions or establishments. Startups and young businesses are inherently engaged in developing products, processes, and customer base but it is likely difficult for such firms to break out separate expenditures. Third, more generally, traditional R&D activity is a narrower concept of innovative activity compared to the broader perspective discussed below on intangible capital accumulation.

Measuring innovative activities using patents is another commonly used alternative approach. Using patents and patent citations as indicators of innovation has a long history (see, the survey by Griliches (1990)). Patents and patent citations as indicators suffer from many of the same limitations as R&D expenditures. They are more informative in some sectors and technologies compared to others. Pavitt (1988) argues that they offer differential protection across sectors and technologies. This also leads to differential propensities across firms to patent their innovative activity. Like R&D expenditures, this means that patents may miss important innovative activity. On the flip side, patenting activity can also provide false positives for innovation when they are used as a defensive measure. Interestingly, GK use patents as one of their three measures of innovation in their empirical exercises, but conclude "that patents are not a good measure of the rate of technological change (p.650)."

There have been research efforts to integrate the R&D and patent data into the LBD. For example, the research by Acemoglu et al. (2017) takes advantage of such integration in a manner closely related to the issues we address in this paper. Specifically, they find that in the innovation intensive industries (essentially industries with sufficient R&D and patent activity) young firms are the most innovation intensive as measured by innovation to sales ratios. Their analysis shows the potential promise of such data integration.[12] However, Acemoglu et al. (2017) focus on only about five percent of all U.S. firms.

Another example is the ongoing Census Bureau project integrating measures of innovation into the LBD to enhance both the Business Dynamic Statistics and the data

[12] This research predates the development of the revenue enhanced LBD, which is what we use in the empirical part of this paper.

infrastructure available to the research community (through the Federal Statistical Research Data Center network). The strategy is to produce an indicator for innovation based upon a multi-dimensional concept that can encompass measures such as R&D expenditures and patents (as well as indicators such as being part of an industry that is high tech, see Goldschlag and Miranda (2016)). One of the first steps in this research project is building a firm-level indicator of patenting activity.

Building upon the experience of earlier researchers linking patent activity to the LBD, Graham et al. (2015) supplement this with linked employee-employer data from the Longitudinal Employer Household Dynamics (LEHD) infrastructure.[13] This allows them to link not only on the business assignee name but also on inventor names listed increasing their match rate to about 90% an improvement over earlier other efforts of about 70-80%. Improvements in the matching and imputation methodologies to create an integrated data infrastructure are currently areas of active research; future research may delve into some measure of patent quality.

An ongoing challenge for any attempt to measure innovation is sparsity. In any dataset, such rare behavior shows up as possibly long gaps in incumbent innovator activity. Alternatively, it also implies that most business start-ups do not engage in traditionally defined innovative activity. Hurst and Pugsley (2011) find that "most surviving small businesses do not innovate along any observable margin. Very few report spending resources on research and development, getting a patent, or even obtaining copyright or trademark protection for something related to the business.(p.74)." This finding that many startups are not inclined towards being innovative but are instead "lifestyle" entrepreneurs, is not inconsistent with the literature that finds that startups are an important source of innovation. As Acemoglu et al. (2017) note, startups and young firms are more innovative than older firms but this is conditional on the startups and young firms being in innovative intensive industries. Similarly, Graham et al. (2015) find that patenting is a relatively rare event for small firms but that most patenting firms are small. It also points to the importance of taking into account innovative activity not well captured by traditional measures.

[13] The LEHD program has worker level information matched to businesses for much of the private employers in the U.S. The core of LEHD data are wage records from State Unemployment Insurance programs linked to establishment data from the Quarterly Census of Employment and Wages (QCEW). The number of records in LEHD data has increased over time as states have joined the voluntary partnership; in the most recent year, the LEHD data tracks more than 130 million worker records each quarter.

Recognizing the potential importance of these innovative activities for startups and young businesses, the Annual Survey of Entrepreneurs (ASE) included an innovation module for 2014 and adjusted the sample to try to capture innovative firms (see Foster and Norman (forthcoming)).[14] The ASE module on innovation is based upon parts of NSF's Microbusiness Innovation Science and Technology (MIST) Survey. The ASE innovation module has eight questions combining questions on inputs to innovation and direct measures of innovation. For the former, information is collected on the types R&D activities, their cost and funding, and the associated number of employees engaged in R&D. Direct measures of innovation are in questions concerning process and product innovations where innovations are broadly defined to include products or processes new to the market and those new to the firm. Process innovation questions focus on the nature of the innovation such as whether it is a new way to make purchases or a new way to deliver goods or services. Furthermore, in recognition of the importance of these small firms in innovation, the Census Bureau fielded in 2017 a version of the BRDIS that targets micro businesses (Business R&D and Innovation for Microbusinesses Survey). Our analysis suggests that integrating the ASE with the revenue enhanced LBD has considerable promise.

Starting with reference year 2017, the ASE will be subsumed into the Annual Business Survey (ABS). In partnership with NSF, the Census Bureau intends to conduct the ABS for reference years 2017-2021. This firm-level survey will also replace the Survey of Business Owners and Business R&D and Innovation for Microbusinesses Survey (however, the Business R&D and Innovation Survey will continue to collect R&D activities for firms with 10 or more employees). Thus the ABS will bring together four areas of inquiry: business characteristics (including financing); owner characteristics (including gender, ethnicity, race, and veteran status); research and development activity and costs (for small firms only); and innovation activities. The ABS is also planned to have a modular component for topical questions. Approximately 850,000 employer firms will be sampled in the baseline reference year of 2017 and 300,000 in the remaining years (see Foster and Norman (forthcoming)).

[14] The ASE will produce annual data on economic and demographic characteristics of employer businesses and their owners by gender, ethnicity, race, and veteran status for reference years 2014-2016. The ASE represents a partnership between the Census Bureau, the Ewing Marion Kauffman Foundation and the Minority Business Development Agency (MBDA).

With fully integrated data, the type of analysis conducted in this paper could be greatly enhanced. Such analysis would permit direct measures of innovation that would be useful for either hypothesis testing or external validity checks. Direct measures integrated with other elements of the data infrastructure would be very useful also for exploring possible heterogeneity in the type of dynamics we have discussed.

In addition, our findings suggest that tracking the joint dynamics of entry, productivity dispersion, and productivity growth offers a potentially useful cross check for traditional measures of innovation. Suppose, for example, that we observe GK dynamics in an industry where the traditional measures of R&D and patents do not capture innovation, this could suggest that this is an industry where these traditional measures are less informative about innovation dynamics. We regard combining the indirect and direct approaches to measuring innovative activity as a high priority for future research.

5.2 Linking Entry and Innovation

Our analysis suggests a tight link between surges in innovation and entry; however, there are open "chicken and egg" questions about their respective timing and interactions. For example, a surge of innovation may occur first at incumbent firms and this could induce entry. Alternatively, the surge of innovation may occur jointly with the surge in entry because innovators create new firms to engage in innovative activity. The GK model distinguishes between these two sources and their impacts: innovations from incumbent firms tend to produce incremental changes, innovations from sources outside the set of current producers tend to produce transformational changes and thus induce entry. While some evidence and models suggest the latter is important (see, Acemoglu et al. (2017)), it is possible that the dynamics are more subtle so this remains an open area of measurement and research.

One way to investigate this would be to track the career paths of individual innovators and their links to firms. Using the LEHD infrastructure to link the individual innovators into the revenue enhanced LBD would enable exploring the inherent chicken-egg issues about innovation and entry. That is, one could examine whether transformational innovations arise from employees of incumbent firms who then go on to spin-off new firms. If this is the case, it may appear that the innovation occurred outside the incumbent firm when in fact, it was incubated at the incumbent firm.

A challenge here is that innovators may go from being employees of incumbent firms to business owners of new firms and ultimately become employees of the new firm if and when the firm incorporates. This implies that tracking the career history of innovators will also involve tracking business owners. Administrative and survey data on business owners will thus need to be integrated into the data infrastructure. A team at Census is exploring the use of person-level business owner identifiers in the administrative data for this purpose.[15] Our analysis highlights the substantial payoffs from such data integration as this has the potential to greatly enhance our understanding of the connection between entrepreneurship and innovation as well as the subsequent productivity and job growth gains from such activity.

A related challenge here is that, as discussed above, there are a host of factors that impact the incentives for entry into entrepreneurship that may yield surges or declines in the pace of entry within a country or industry. Recall that the Hurst and Pugsley (2011, 2014) hypothesis is that non-pecuniary benefits is the primary driver of many entrepreneurs and that entrepreneurs entering for such motives are concentrated in specific industries. A surge or decline in entry in such sectors may have little to do with innovation and productivity growth. Sorting the relative importance of the factors influencing entrepreneurship is an important area of research. The ASE includes some questions concerning the motivations and aspirations of entrepreneurs (see Foster and Norman (forthcoming) for a discussion). One question concerns reasons for owning the business and includes checkboxes for responses such as: “Wanted to be my own boss” and “Best avenue for my ideas/goods services.” A second question asks about aspirations for the business over the next five years and includes checkboxes for responses such as: “Larger in terms of sales or profits” and “About the same amount of sales or profits.” Tracking the career history of entrepreneurs as well as incorporating information about the activities and incentives of entrepreneurs as discussed in this section should help in this effort.

5.3 Intangible Capital

Another interesting area of inquiry is to relate the innovative activities associated with entrants and young firms to the growing literature on measuring and understanding the growth of intangible capital. One interpretation of our work in this paper is that we use entry as a proxy for

[15] A recent example of the application of integrated data is see Bell et al. (2017). Using data on patents, tax records and test scores from New York school districts they show that family income and exposure to innovative during childhood significantly increase the propensity to become inventors.

innovation. It might be fruitful to think about the time and resources associated with entry and young firm activity as a measure of intangible capital investment. This perspective is consistent with the broad view of intangible capital of Corrado, Hulten, and Sichel (2005, 2009) and Corrado et al. (2013) who define intangible capital expenditures as any current period expenditures by firms intended to enhance future production or sales. Other studies apply narrower definitions of innovation and intangible investment focusing on the effects of spending on specific categories of intangible assets, such as employer funded training, software, R&D, branding and design, and process improvement (see Awano et al. (2010) for more details). A recent example of estimating the contribution of innovation and intangible investment to growth can be found in Haskel et al. (2014). Exploring such issues within the context of the joint dynamics of entry, dispersion, and growth would be of considerable interest.

We think a strong case can be made that entrants and young firms are inherently engaged in intangible capital investment. Likewise young firms are engaged in activity to develop products and processes and to break into markets (such as developing a customer base – see Foster, Haltiwanger, and Syverson (2016)). The experimentation phase we have discussed, and provided some evidence in support of, is another form of investment in activity. Kerr et al. (2014) make a related point in arguing that "entrepreneurship is fundamentally about experimentation."

Exploring how to measure and track the indirect approach we have advocated in this paper within the context of the measurement and contribution of intangible capital would be of considerable interest. Haltiwanger, Haskel and Robb (2010) discuss and consider some promising possibilities for tracking intangible capital investment by new and young firms. For example, they find that young firms appear to be actively investing in various forms of intangible capital (using tabulations from the Kauffman Firm Survey that query firms about their activities). Even though they find supporting evidence, they highlight the difficulties of obtaining such measures from entrants and young firms. The founders and employees of new firms are engaged in many tasks so that probing questions are needed to elicit the time and resources that should be considered intangible capital investment.

Overall, our view is that the conceptual approach of intangible capital investment advocated by Corrado, Hulten and Sichel (2005, 2009) has the greatest potential for direct measurement of investment in innovative activities. This approach takes the appropriate broad based perspective on innovative investment activities and advocates capitalizing expenditures on these activities in the same manner as for physical capital expenditures. The challenge here is developing measurement instruments that can capture such intangible investment activities for all sectors and all firms including young and small firms. We regard our indirect approach as complementary to the intangible capital approach. Combined with suitable measures of the latter, our approach could be used to study the stochastic and uncertain payoffs of investments in intangible capital. In addition, as with direct approaches to measure innovation, our indirect approach could be used for external validity or cross validation of intangible capital measurement.

5.4 High Versus Low Frequency Variation

Understanding high versus low frequency productivity dispersion and productivity growth dynamics would be another useful area of inquiry. The empirical results in Section 4 suggest that an increase in industry-specific entry rates leads to increases first in productivity dispersion and then productivity growth. As emphasized above, we estimated these relationships using low frequency variation with the express intent of abstracting from cyclical dynamics. Since the contribution of innovation may materialize with potentially long and variable lags (see Griliches (1984)) and may even arrive in multiple waves (Syverson (2013)), long-run variation seems more appropriate for estimation purposes.

On the other hand, the appropriate horizon at which other factors affect dispersion is less clear a priori. Some of the results in the literature on frictions and distortions are based on annual average indicators (see for example, Hsieh-Klenow (2009) or Foster et al. (2017)). In addition, other evidence indicates that the effect of changing frictions may also be detected at higher frequencies. A recent example is Brown et al. (2016), who find that yearly dispersion measures increase during and after periods of deregulation.

While it may be of interest to abstract from short-run variation for certain research questions, it may be that cyclical dynamics are present and interact with lower frequency dynamics. For example, Kehrig (2015) and Bloom (2009) document that within-industry productivity dispersion varies negatively with the cycle: it is greater in recessions than in booms. In addition, there is evidence that periods of Schumpeterian creative destruction coincide with recessions – although the extent to which this holds varies over cycles (see, e.g., Foster, Grim, and Haltiwanger (2016)).

To help illustrate the effects of these complicating factors we have estimated simple two-variable panel vector autoregression specifications (VAR) using annual time series on productivity dispersion, entry, and productivity growth from pooling High Tech industries (4-digit NAICS) between 1997 and 2013. One can think of these VARs as high frequency analogues of the analysis in Section 4. We show these high frequency results to illustrate that developing a way to think about low and high frequency dynamics in an integrated manner would be of interest.

All the results reported below are derived from stable first order VARs, where the underlying coefficients and standard errors are GMM-based and the lag order of the VAR is implied by standard information criteria.[16] The first impulse response function is estimated using changes in entry and dispersion, with this Cholesky ordering. Results are shown in Panel A of Figure 6: dispersion increases significantly in the wake of a positive change in entry, and the effect lasts 2-3 years. This finding is broadly consistent with our findings in Section 4.[17]

However, investigation of other high frequency dynamics reminds us of the many different factors in the joint distribution of entry, productivity dispersion and productivity growth. Using a two variable VAR relating productivity dispersion and productivity growth, we find (in unreported results) evidence of Granger causality from productivity growth to productivity dispersion. Panel B of Figure 6 shows that a positive (high frequency) productivity shock has a short-lived negative response on within-industry productivity dispersion. The short-lived negative response may be consistent with a number of theories. First, it may reflect the effect of cyclical variation in uncertainty (see Bloom (2009)). Alternatively, a negative response may be related to demand-driven fluctuations in the price of fixed inputs that lead to positive selection among more productive firms (see Kehrig (2015) for more details). There are other possibilities as well as there might be some interaction between these high frequency dynamics and the lower frequency dynamics that have been the focus of much of our discussion in this paper.

[16] We use the Stata module documented in Abrigo and Love (2016). The module integrates all the necessary steps of the empirical implementation: parameter estimation, hypothesis testing, lag order selection and impulse response estimation.

[17] The two variables in this VAR are the industry entry rate and the change in the within-industry productivity dispersion. The panel VAR has industry effects but no year effects. This is different from our low frequency panel regressions in Section 4 which relate entry to the level of within industry dispersion controlling for both industry and time effects. We used forward orthogonal deviations to remove industry effects from the VARs because this transformation tends to outperform the first-difference transformation when using GMM, see Hayakawa (2009).

A potentially useful approach would be to investigate the empirical performance of cointegrating relationships. The main advantage of the concept of cointegration would be its straightforward use in decomposing time series variation into high- and low-frequency dynamics, especially if it is reasonable to assume that different forces generate variation at different frequencies. For example, when the long-run dynamics of entry and dispersion are related because of innovative activity, and short-run variation is associated with business cycle fluctuations.

In short, this discussion of high versus low frequency variation highlights some of the challenges and limitations of the indirect approach we are advocating in this paper. While we think much can be learned from our indirect approach, many other factors impact the joint dynamics of entry, dispersion and productivity growth that need to be considered and controlled for in trying to draw inferences about innovative activity.

6. Concluding Remarks

Our findings suggest there are rich joint dynamics of firm entry, within-industry productivity dispersion across firms, and within-industry productivity growth that help shed light on innovation dynamics. The patterns are broadly consistent with models of innovation where periods of rapid innovation are accompanied by a surge of entry. Such a surge in entry induces a rise in productivity dispersion and productivity growth within industries. Productivity growth stems from within-firm productivity growth by, and reallocation of resources towards, successful innovators.

Our analysis is intended to be exploratory. Our objective is to discuss the conceptual and measurement challenges exploring these joint dynamics that appear to be important for understanding the complex nature of innovation. Part of the conceptual challenge is that alternative factors may influence the joint dynamics of entry, productivity dispersion, and productivity growth. For example, changes in frictions and distortions yield a distinct pattern of co-movement as emphasized by the recent literature. Second moment shocks to the distribution of idiosyncratic productivity and profitability at cyclical or other frequencies are also likely important.

In terms of measuring innovation, there are efforts underway (and already interesting research based on such efforts) to integrate traditional measures such as patents, R&D

expenditures, and the indicators of firm and industry dynamics that are a focus of our analysis. We think there will be substantial payoff from such efforts at further data integration. We also emphasize that even as this effort becomes increasingly realized, some open questions will remain. For example, if we can detect the presence of an innovative period in an industry as suggested by the results of this paper then it will be interesting to cross-check the joint dynamics of entry, productivity dispersion, and productivity growth against traditional innovation measures.

We also think that the indirect approach suggested in this paper can be fruitfully combined with the ongoing efforts to measure innovative activity by capitalizing intangible capital expenditures. Combining these efforts would serve as a useful cross check but also provide the ability to investigate the stochastic and heterogeneous returns to investment in intangible capital expenditures. Combining these efforts would also help overcome the known challenge that traditional indicators such as R&D expenditures and patents may not capture the full extent of young firms' innovative activity. Capturing intangible investment by entrants and young firms is especially challenging since the founders and workers at young firms are inherently engaged in multi-tasking as they try to survive and ramp up production and their customer base for the future.

It is our view that overcoming these conceptual and measurement challenges will involve a multi-dimensional approach. First, is continuing and expanding the integration of both person-level and business-level data. Currently, these include both survey and administrative sources, but they could also include commercial data. Second is continuing efforts to link these data longitudinally and to improve these links. Third, is using a more focused approach to survey content; to use special modules like in Annual Survey of Entrepreneurs (or the forthcoming Annual Business Survey) to ask deeper questions about hard-to-measure concepts such as intent to innovate. Fourth, is using economic relationships between relatively easy-to-measure concepts (such as entry and productivity dispersion) to help direct our measurement efforts towards areas of the economy where innovation is taking place. The payoff from these efforts could be substantial. It will only be through such efforts that we can understand the complex and noisy process through which innovation leads to productivity and job growth.

References

Abrigo, Michael R.M. and Inessa Love, 2016. "Estimation of Panel Vector Autoregression in Stata: a Package of Programs," Working Papers 201602, University of Hawaii at Manoa, Department of Economics.

Acemoglu, Daron, Ufuk Akcigit, Harun Alp, Nicholas Bloom, and William R. Kerr. 2017. "Innovation, Reallocation, and Growth." NBER Working Paper No. 18993.

Andrews, Dan, Chiara Criscuolo, and Peter N. Gal. 2016. "Frontier Firms, Technology Diffusion and Public Policy: Micro Evidence from OECD Countries." OECD: *The Future of Productivity: Main Background Papers.*

Awano, Gaganan, Mark Franklin, Jonathan Haskel, and Zafeira Kastrinaki, 2010. "Measuring investment in intangible assets in the UK: results from a new survey*," Economic & Labour Market Review*, Palgrave Macmillan; Office for National Statistics, 4(7): 66-71.

Bartelsman, Eric, John Haltiwanger and Stefano Scarpetta. 2013. "Cross-Country Differences in Productivity: The Role of Allocation and Selection." *American Economic Review*, 103(1): 305-34.

Baily, Martin N., Eric J. Bartlesman, and John Haltiwanger. 2001. "Labor Productivity: Structural Change and Cyclical Dynamics," *The Review of Economics and Statistics* 83(3): 420-433.

Baily, Martin N., Charles Hulten, and David Campbell. 1992. "Productivity Dynamics in Manufacturing Plants," *Brookings Papers on Economic Activity.* Microeconomics, 187-267.

Bell, M Alexander, Raj Chetty, Xavier L. Jaravel, Neviana Petkova, John Van Reenen. 2017. "Who Becomes An Inventor in America? The Importance of Exposure to Innovation" NBER Working Paper No. 24042.

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

Bloom, Nicholas, Erik Brynjolfsson, Lucia Foster, Ron Jarmin, Megha Patnaik, Itay Saporta-Eksten, John Van Reenen, 2018. "What Drives Differences in Management Practices?" *American Economic Review*, forthcoming.

Brown, Lawrence, Thomas Plewes, and Marisa Gerstein. 2005. *Measuring Research and Development Expenditures in the U.S. Economy*, National Academy of Sciences Press.

Brown, J. David, Emin Dinlersoz, John S. Earle, 2016. "Does Higher Productivity Dispersion Imply Greater Misallocation? A Theoretical and Empirical Analysis," CES Working Paper 16-42.

Byrne, David, Dan Sichel, and Marshall Reinsdorf (2016). "Does the United States have a Productivity Slowdown or a Measurement Problem?" *Brookings Papers on Economic Activity*, Conference Draft, March.

Byrne, David, Stephen Oliner, and Dan Sichel (2013). "Is the Information Technology Revolution Over?" *International Productivity Monitor*, Number 25, Spring.

Cooper, Russell W. and John C. Haltiwanger. 2006. "On the Nature of Capital Adjustment Costs," *Review of Economic Studies*, 73(3): 611-633.

Corrado, Carol, Charles Hulten, and Dan Sichel. 2005. "Measuring Capital and Technology: An Expanded Framework." In Measuring Capital in the New Economy, Vol. 65 (eds., Carol Corrado, John Haltiwanger, and Dan Sichel). Chicago: The University of Chicago Press.

Corrado, Carol, Charles Hulten, and Dan Sichel. 2009. "Intangible Capital and US Economic Growth". *The Review of Income and Wealth*, (55:3), pp. 661-685.

Corrado, Carol, Jonathan Haskel, Cecilia Jona-Lasinio, Massimiliano Iommi. 2013. **"**Innovation and Intangible Investment in Europe, Japan and the United States" *Oxford Review of Economic Policy*, 29(2), 261–286.

Cunningham, Cindy, Lucia Foster, Cheryl Grim, John Haltiwanger, Sabrina Wulff Pabilonia, Jay Stewart, and Zoltan Wolf. 2017. "Dispersion in Dispersion: Measuring Establishment-Level Differences in Productivity." Unpublished paper.

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

Decker, Ryan, John Haltiwanger, Ron S. Jarmin, and Javier Miranda. 2016. "The Decline of High Growth Young Firms in the U.S.: Where has all the Skewness Gone?" *European Economic Review*, 86, 4-13.

Decker, Ryan, John Haltiwanger, Ron S. Jarmin, and Javier Miranda. 2017. "Declining Dynamism, Allocative Efficiency, and the Productivity Slowdown" *American Economic Review Papers and Proceedings*, May.

Decker, Ryan, John Haltiwanger, Ron S. Jarmin, and Javier Miranda. 2018. "Changing Business Dynamism: Shocks vs. Responsiveness" NBER Working Paper No. 24236.

Dunne, Timothy, Mark J. Roberts, and Larry Samuelson. 1989. "The Growth and Failure of US Manufacturing Plants. *The Quarterly Journal of Economics*, 104(4), 671-98.

Ericson, Richard and Ariel Pakes. 1995. Markov-Perfect Industry Dynamics: A Framework for Empirical Work. *The Review of Economic Studies* 62, no. 1: 53-82.

Fernald, John. 2014. "Productivity and Potential Output Before, During, and after the Great

Recession." *NBER Macroeconomics Annual* 29: 1-51.

Foster, Lucia and Patrice Norman. Forthcoming. "The Annual Survey of Entrepreneurs." *Journal of Economic and Social Measurement*.

Foster, Lucia, Cheryl Grim, and John Haltiwanger. 2016. "Reallocation in the Great Recession: Cleansing or Not?" *Journal of Labor Economics* 34(S1), Part 2, 293-331.

Foster, Lucia, Cheryl Grim, and Nick Zolas. 2016. "A Portrait of Firms that Invest in R&D." CES Working Paper 16-41.

Foster, Lucia, Cheryl Grim, John Haltiwanger, and Zoltan Wolf. 2016. "Firm-Level Dispersion in Productivity: Is the Devil in the Details?" *American Economic Review Papers and Proceedings* 106(5), 95-98.

Foster, Lucia, Cheryl Grim, John Haltiwanger, and Zoltan Wolf. 2017. "Macro and Micro Dynamics of Productivity: From Devilish Details to Insights." NBER Working Paper No. 23666.

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.

Foster, Lucia, John Haltiwanger, and Chad Syverson. 2016. "The Slow Growth of New Plants" Learning About Demand?" *Economica*, 83(329), 91-129.

Foster, Lucia, John Haltiwanger, and CJ Krizan. 2001. "Aggregate Productivity Growth: Lessons from Microeconomic Evidence," in New Developments in Productivity Analysis (eds. Edward Dean, Michael Harper and Charles Hulten), University of Chicago Press.

Goldschlag, Nathan and Javier Miranda, 2016. "Business Dynamics Statistics of High Tech Industries," CES Working Paper 16-55.

Gordon, Robert J., 2016. "The Rise and Fall of American Growth: The U.S. Standard of Living since the Civil War". Princeton: Princeton University Press.

Gort, Michael and Steven Klepper. 1982. "Time Paths in the Diffusion of Product Innovations," *The Economic Journal*, Vol. 92, No. 367 (Sept.), 630-653.

Graham, Stuart, Cheryl Grim, Tariqul Islam, Alan Marco, and Javier Miranda (2015). "Business Dynamics of Innovating Firms: Linking U.S. Patents with Administrative Data on Workers and Firms," CES Working Paper 15-19.

Griliches, Zvi, 1980. "Returns to Research and Development Expenditures in the Private Sector," in New Developments in Productivity Measurement, John W. Kendrick and Beatrice N. Vaccara, eds., University of Chicago Press.

Griliches, Zvi. 1984. "Introduction to R&D, Patents, and Productivity," in R&D, Patents, and Productivity, Zvi Griliches, ed., University of Chicago Press.

Griliches, Zvi. 1990. "Patent Statistics as Economic Indicators: A Survey." *Journal of Economic Literature*, XXVIII: 1661-1707.

Griliches, Zvi, and Haim Regev. 1992. “Productivity and Firm Turnover in Israeli Industry: 1979-1988,” NBER Working Paper No. 4059.

Haltiwanger, John. 2016. “Firm Dynamics and Productivity: TFPQ, TFPR and Demand Side Factors.” *Economia*, 17(1), 3-26.

Haltiwanger, John, Ian Hathaway and Javier Miranda. 2014. “Declining Business Dynamism in the U.S. High-Technology Sector”, Kauffman Foundation.

Haltiwanger, John, Jonathan Haskel, and Alicia Robb. 2010. “Extending Firm Surveys to Measure Intangibles: Evidence from the United Kingdom and United States.” Unpublished paper.

Haltiwanger, John, Robert Kulick and Chad Syverson. 2018. “Misallocation Measures: The Distortion that Ate the Residual.” NBER Working Paper No. 24199.

Haltiwanger, John, Ron S. Jarmin, Robert Kulick, and Javier Miranda. 2017. “High Growth Young Firms: Contribution to Job, Output and Productivity Growth” in Measuring Entrepreneurial Businesses: Current Knowledge and Challenges, NBER/CRIW Conference Volume published by University of Chicago Press.

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

Haskel, Jonathan, Peter Goodridge, Gavin Wallis. 2014. “UK investment in intangible assets: Report for NESTA,” Working Papers 12846, Imperial College, London, Imperial College Business School.

Hayakawa, Kazuhiko 2009. “First Difference or Forward Orthogonal Deviation- Which Transformation Should be Used in Dynamic Panel Data Models?: A Simulation Study,” Economics Bulletin, AccessEcon, 29(3), 2008-2017.

Hecker, Daniel E. 2005. “High-technology employment: a NAICS-based update.” *Monthly Labor Review*, 128: 57.

Hopenhayn, Hugo. 1992. “Entry, Exit, and Firm Dynamics in Long Run Equilibrium.” *Econometrica*, 60(5): 1127-50.

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

Hottman, Colin J., Stephen J. Redding, and David E. Weinstein, 2016. “Quantifying the Sources of Firm Heterogeneity” *Quarterly Journal of Economics*, 131(3): 1291-1364.

Hsieh, Chang-Tai and Peter J. Klenow. 2009. “Misallocation and Manufacturing TFP in China and India,” *Quarterly Journal of Economics*, 124(4):1403-1448.

Hurst, Erik and Benjamin Pugsley, 2011. “What Do Small Businesses Do?” *Brookings Papers on Economic Activity*. 43(2): 73-142.

Hurst, Erik G., and Benjamin Pugsley 2017. “Wealth, Tastes, and Entrepreneurial Choice.” in Measuring Entrepreneurial Businesses: Current Knowledge and Challenges, NBER/CRIW Conference Volume published by University of Chicago Press.

Jarmin, Ron S., and Javier Miranda. 2002. “The Longitudinal Business Database.” CES Working Paper, 02-17.

Jovanovic, Boyan. 1982. "Selection and the Evolution of Industry." *Econometrica* 50 (May): 649-70.

Kehrig, Matthias, 2015. “The Cyclicality of Productivity Dispersion,’ CES Working Paper 11-15 Revised 2015.

Kerr, William R. and Ramana Nanda and Matthew Rhodes-Kropf. 2014. “Entrepreneurship as Experimentation,” *The Journal of Economic Perspectives*, 28(3): 25-48

Klette, Tor and Samuel Kortum. 2004. “Innovating Firms and Aggregate Innovation*,” Journal of Political Economy*, 112(5): 986-1018.

Lentz, Rasmus and Dale T. Mortensen. 2008. “An Empirical Model of Growth Through Product Innovation*,” Econometrica*, 76(6): 1317-73.

Lucas, Robert E. Jr., 1978. “On the Size Distribution of Business Firms,” *Bell Journal of Economics*, The RAND Corporation, 9(2): 508-523.

Melitz, Marc J. 2003. “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity.” *Econometrica*, 71(6): 1695-1725.

Melitz, Marc J. and Sašo Polanec. 2015. “Dynamic Olley-Pakes Productivity Decomposition with Entry and Exit,” *RAND Journal of Economics* 46 (2): 362-375.

Pavitt, Kevin. 1988. “Uses and abuses of patent statistics” in Handbook of Quantitative Studies of Science and Technology, pages 509-536. Ed: A. F. J. van Raan. Elsevier.

Petrin, Amil, T. Kirk White, and Jerome Reiter. 2011. “The Impact of Plant-level Resource Reallocations and Technical Progress on U.S. Macroeconomic Growth.” *Review of Economic Dynamics*, 14(1):3-26.

Restuccia, Diego and Richard Rogerson. 2008. “Policy Distortions and Aggregate Productivity with Heterogeneous Plants,” *Review of Economic Dynamics*, 11(4): 707-720.

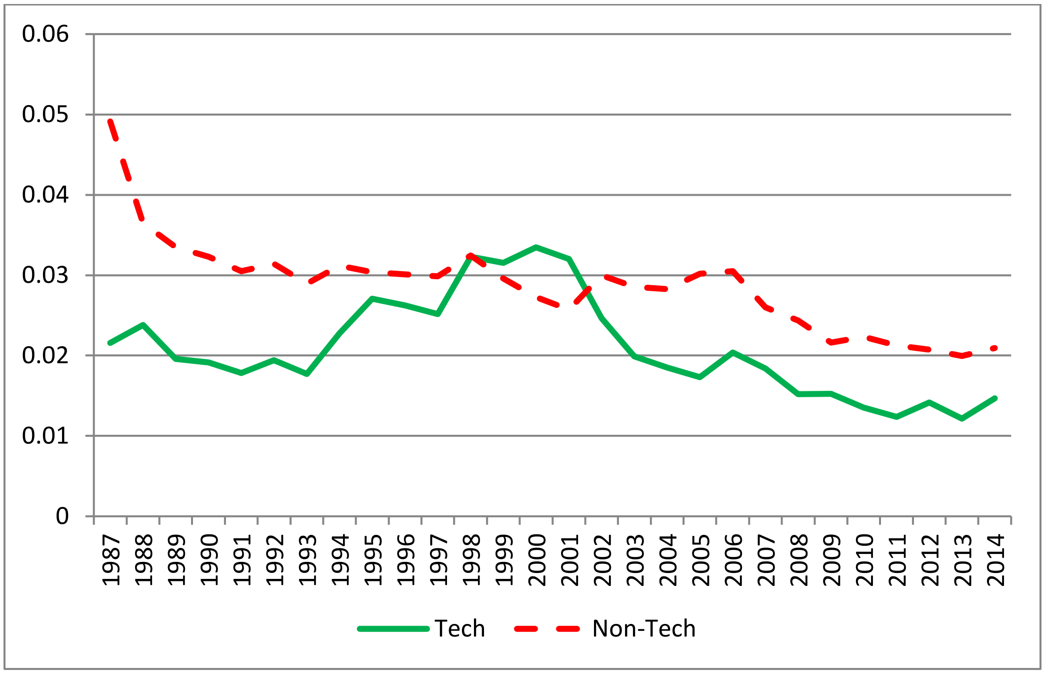
Syverson, Chad. 2013. "Will History Repeat Itself? Comments on "Is the Information Technology Revolution Over*?" International Productivity Monitor*, (25): 37-40

Syverson, Chad. 2011. "What Determines Productivity?" *Journal of Economic Literature*, 49(2): 326-365.

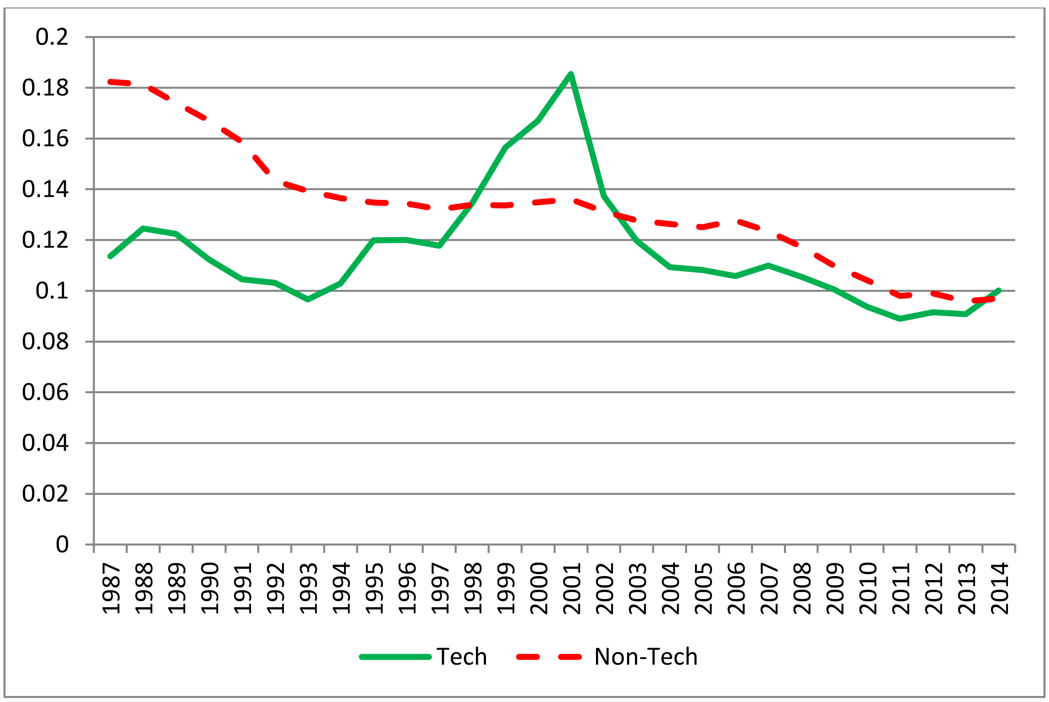
Syverson, Chad. 2004. "Market Structure and Productivity: A Concrete Example." *Journal of Political Economy*, 112(6): 1181-222.

Figure 1: Share of Employment at High Tech and Non Tech

Panel A: Startups

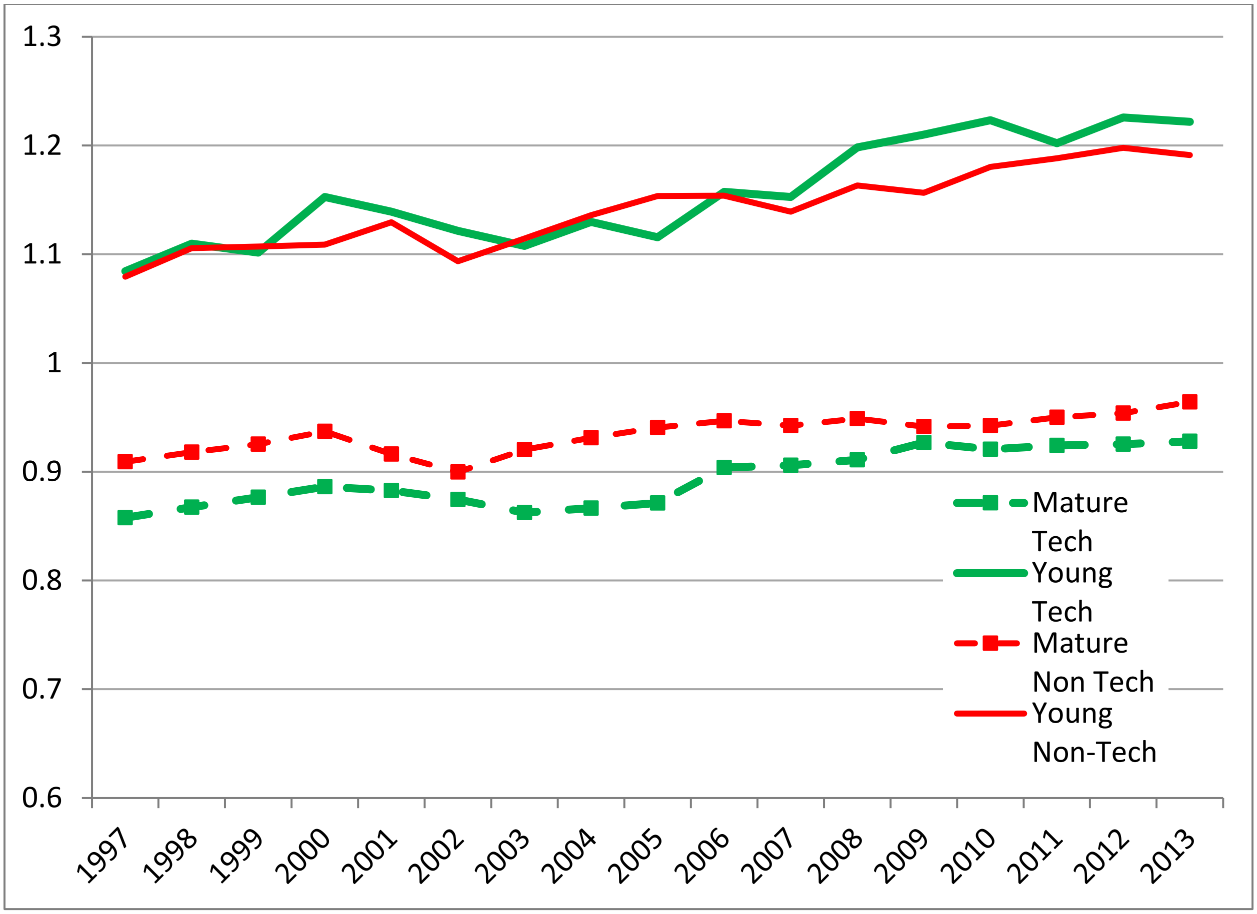


Panel B: Young Firms (Age<5)



Source: Tabulations from the LBD.

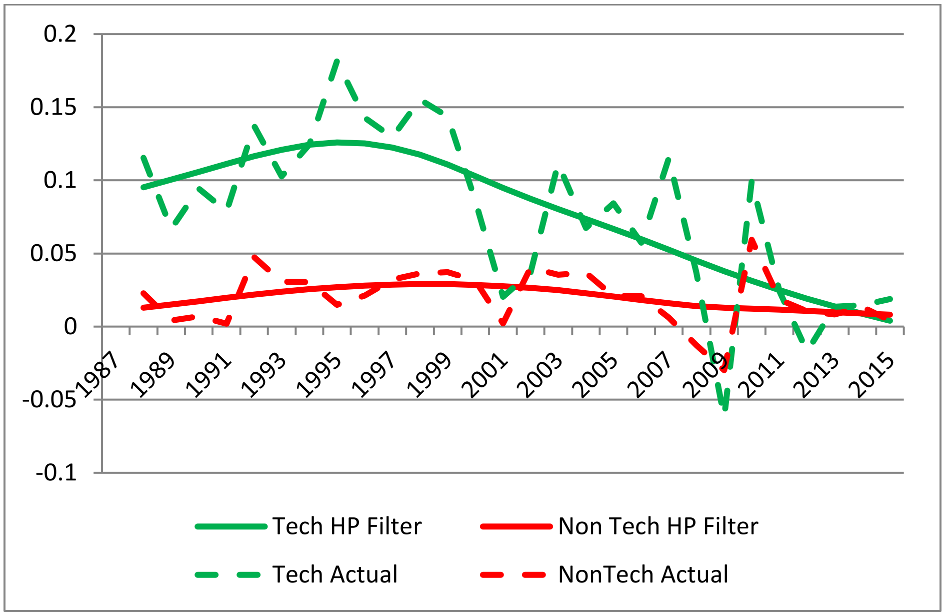
Figure 2: Within-Industry Dispersion in Labor Productivity



Source: Tabulations from the LBD. Dispersion is the inter-quartile range of within industry log revenue per worker. Industry defined at the 4-digit NAICS level.

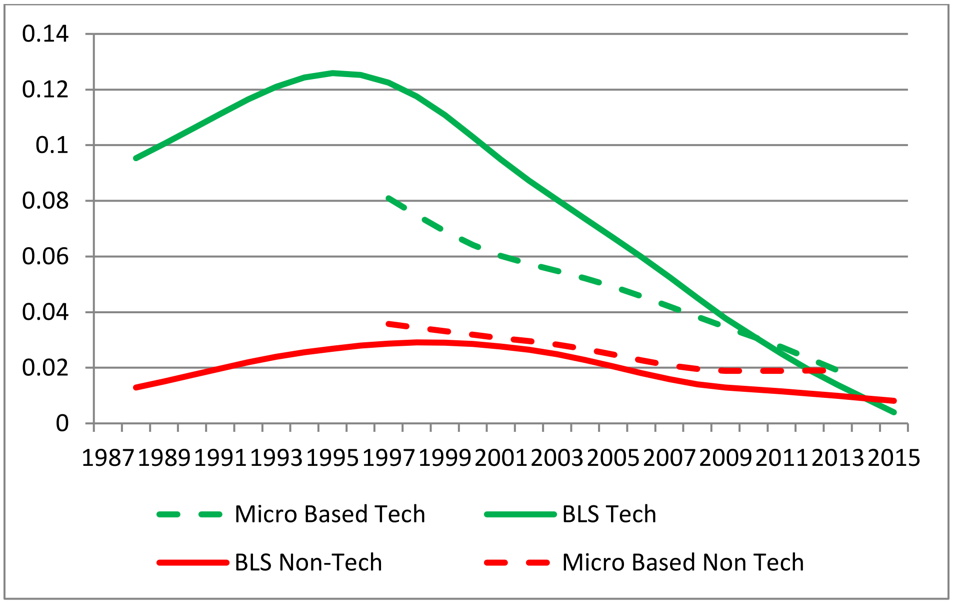
Figure 3: Labor Productivity Growth for High Tech and Non Tech Industries

Panel A: BLS Data Annual and HP Filter



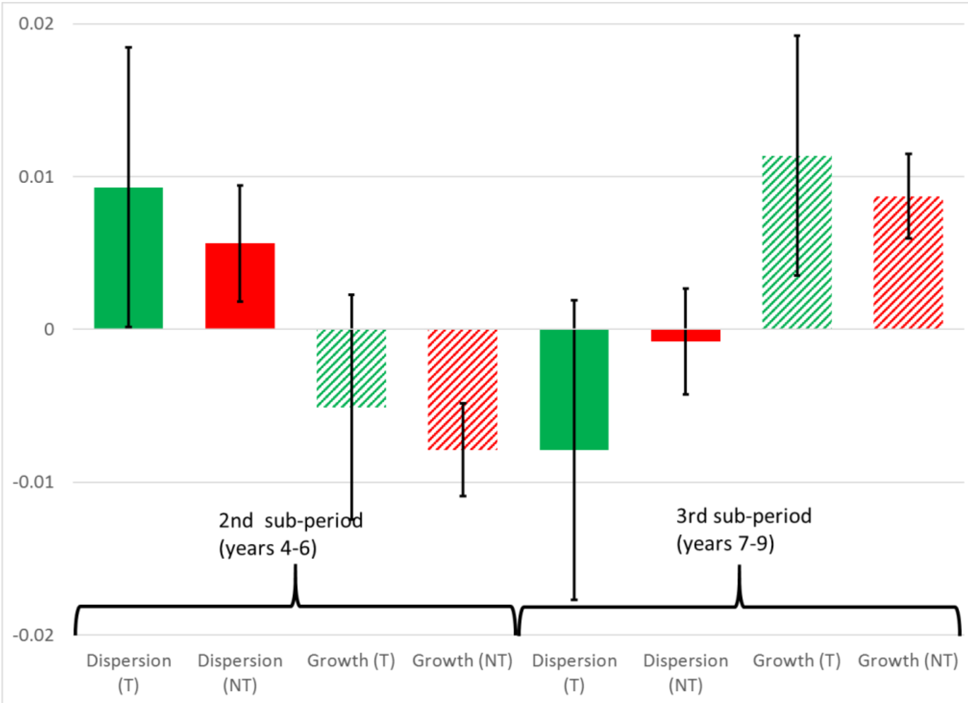
Source: BLS

Panel B: BLS Aggregate Data and Census Micro Data (HP Filtered)



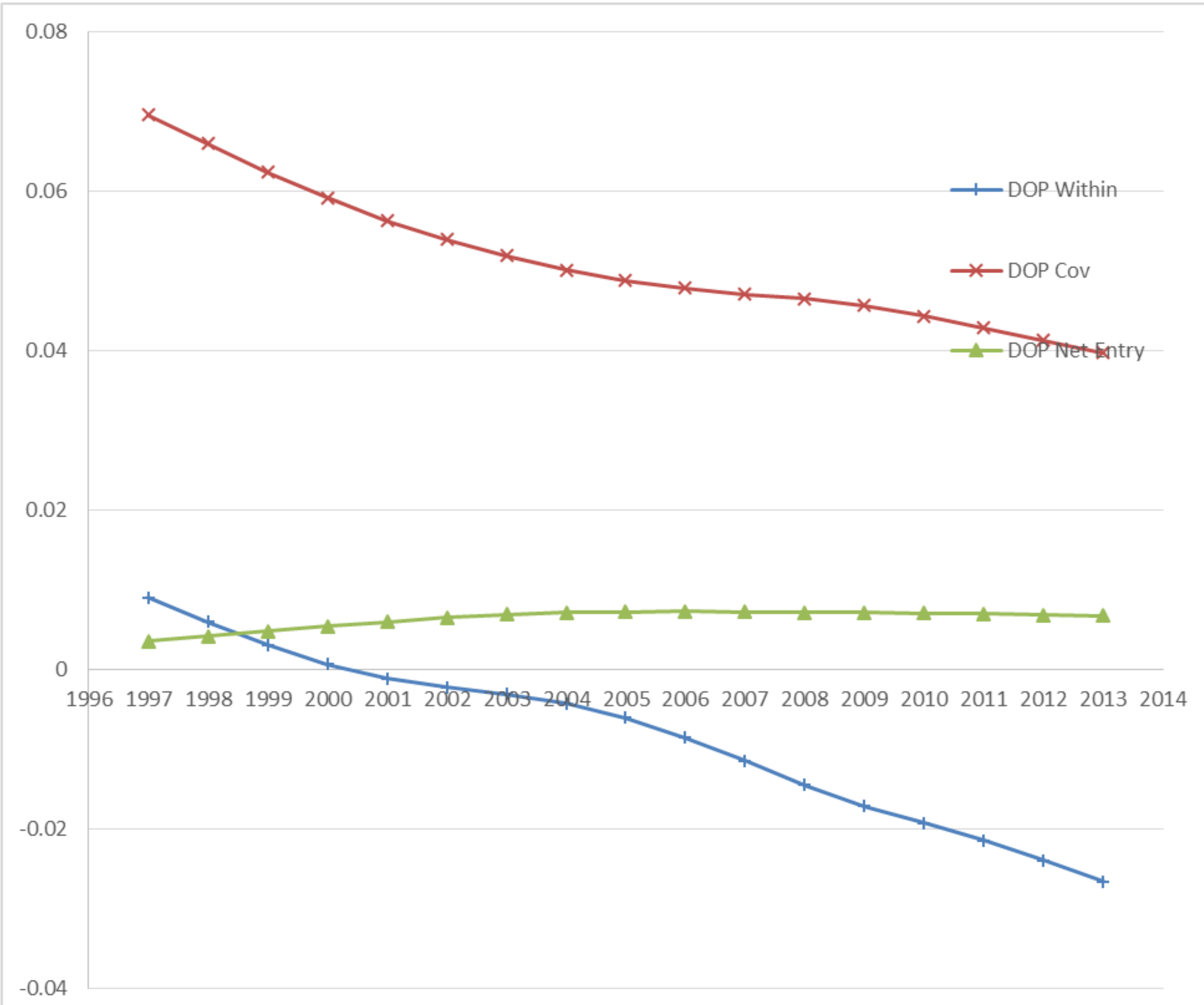
Source: BLS and Tabulations from the dataset described in Haltiwanger, Jarmin, Kulick, and Miranda (2017).

Figure 4: Changes in Productivity Dispersion and Growth from a 1% (one time) Increase in Entry Rate



Source: Authors tabulations from estimated coefficients, see Tables 1 and 2. The left side of the chart represents one sub-period after entry (years 4-6); the right side represents two sub-periods after entry (years 7-9). High Tech (T) results are in green, Non Tech (NT) are in red. Solid bars show results for productivity dispersion; hashed bars show results for productivity growth. The black "whiskers" show approximate 95% confidence intervals. Thus, the first green bar shows the change in 3-year-average productivity dispersion after a 1 percentage point increase in the 3-year-average entry rate for High Tech. All other bars are analogous.

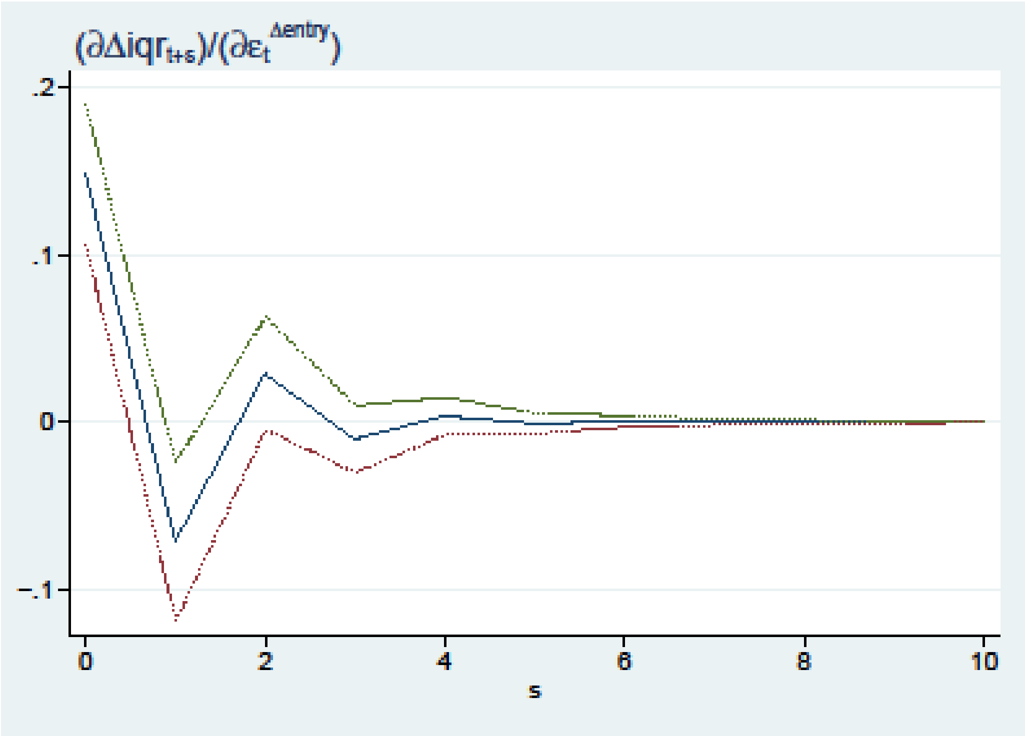
Figure 5: Dynamic Olley-Pakes Decomposition of Aggregate Productivity Growth and Weighted Within-Plant Growth in High Tech industries



Source: Tabulations from the LBD. Decompositions at the 4-digit level for industries in the High Tech sector. 4-digit decomposition averaged across industries using time invariant employment weights.

Figure 6: Impulse Response Functions of Dispersion for High-Tech Sector

Panel A: Positive Entry Shock. Dashed lines show simulated 95% confidence bands.



Panel B: Positive Productivity Shock. Dashed lines show simulated 95% confidence bands

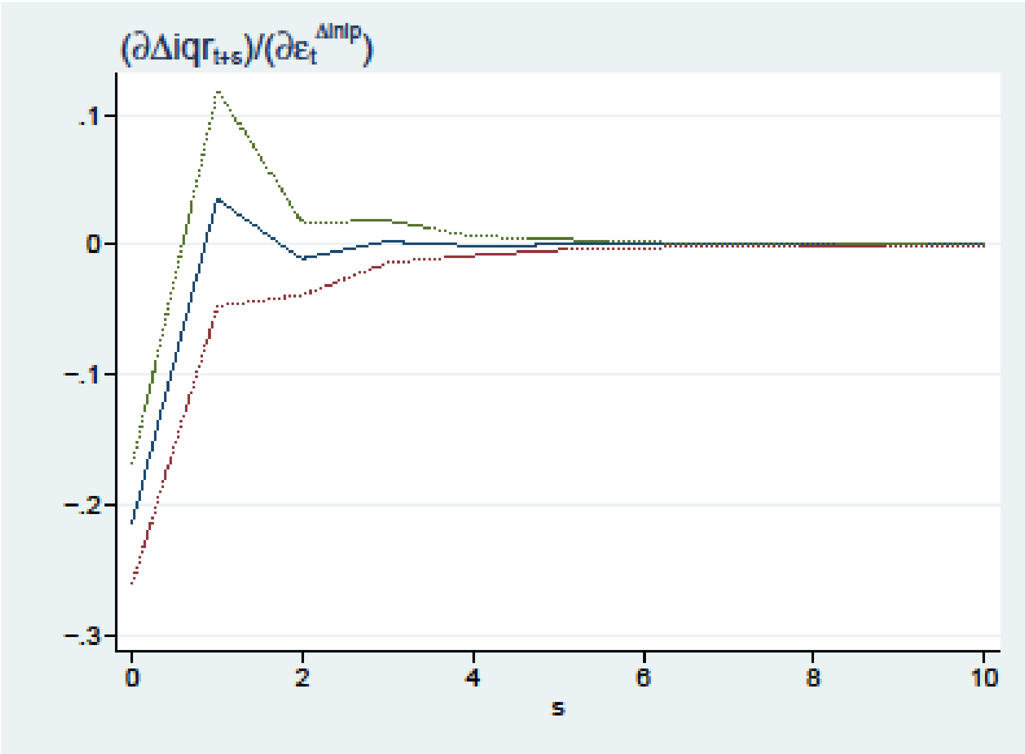


Table 1: Productivity *Dispersion* and Entry	
Lag 1 Entry*Tech	0.929**
	(0.458)
Lag 1 Entry*Non Tech	0.563***
	(0.190)
Lag 2 Entry*Tech	-0.791
	(0.491)
Lag 2 Entry*Non Tech	-0.082
	(0.174)
Industry Effects	Yes
Period Effects	Yes
R-squared	0.93
Number of Observations	1541

Source: Panel regression estimated from industry by year moments computed from the revenue enhanced LBD.

Table 2: Productivity *Growth* and Entry	
Lag 1 Entry*Tech	-0.516
	(0.367)
Lag 1 Entry*Non Tech	-0.791***
	(0.152)
Lag 2 Entry*Tech	1.136***
	(0.393)
Lag 2 Entry*Non Tech	0.871***
	(0.139)
Industry Effects	Yes
Period Effects	Yes
R-squared	0.38
Number of Observations	1541

Source: Panel regression estimated from industry by year moments computed from the revenue enhanced LBD.