Business Dynamics Statistics of High Tech Industries*

Nathan Goldschlag† Javier Miranda‡

March 20, 2018

Abstract

Modern market economies are characterized by the reallocation of resources from less productive, less valuable activities to more productive, more valuable ones. Businesses in the High Tech sector play a particularly important role in this reallocation by introducing new products and services that impact the entire economy. In this paper we describe an extension to the Census Bureau's Business Dynamics Statistics (BDS) that tracks job creation, job destruction, startups, and exits by firm and establishment characteristics including sector, firm age, and firm size in the High Tech sector. We preview the resulting statistics, showing the structural shifts in the High Tech sector over the past 30 years including the surge of entry and young firm activity in the 1990s that reversed abruptly in the early 2000s.

Keywords: firm dynamics, High Tech, startups, employment flows

*Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed. We thank Emin Dinlersoz, John Haltiwanger, Lars Vilhuber and participants at the Center for Economic Studies Seminar Series for helpful comments. All opinions and views expressed are those of the authors and do not represent those of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed.

†U.S. Census Bureau (corresponding author) `nathan.goldschlag@census.gov`

‡U.S. Census Bureau `javier.miranda@census.gov`

Modern market economies are characterized by the reallocation of resources from less productive, less valuable activities to more productive, more valuable ones. Businesses in the High Tech sector play a particularly important role in this reallocation by introducing new products and services that impact the entire economy. In this paper we describe an extension to the Census Bureau's Business Dynamics Statistics (BDS) that tracks job creation, job destruction, startups, and exits by firm and establishment characteristics including sector, firm age, and firm size in the High Tech sector. We preview the resulting statistics, showing the structural shifts in the High Tech sector over the past 30 years including the surge of entry and young firm activity in the 1990s that reversed abruptly in the early 2000s.

1 Introduction

Modern market economies are characterized by the reallocation of resources from less productive, less valuable activities to more productive more valuable ones. Businesses that innovate grow and replace those that do not (Schumpeter, 1942). The innovations that underlie this reallocation are ultimately responsible for increased productivity growth and increasing standards of living. Businesses in the High Tech sector play a particularly important role in this regard, as they bring about new products and services that fuel this reallocation. Not surprisingly, developing a deeper understanding of these businesses, their inception, early life cycle dynamics, and constraints to growth are the focus of intense attention (Stoneman and Battisti, 2010; Acemoglu, Akcigit, Bloom, and Kerr, 2013; Decker, Haltiwanger, Jarmin, and Miranda, 2016a,b). However, our understanding of the dynamics of innovative firms is very limited. One reason for this is that, until recently, longitudinal datasets of establishments and firms covering the U.S. economy did not exist. The development and introduction of the U.S. Census Bureau's Longitudinal Business Database (LBD) changes that. The LBD tracks all establishments and firms in the private non-farm economy with paid employees starting in 1976 and is updated annually.[1] The LBD, however, is a confidential micro dataset only

[1]See Jarmin and Miranda (2002) and Stinson, White, and Lawrence (2017) for details

accessible to Census Bureau employees and researchers with an approved "need to know" and therefore access to the data is necessarily limited. In this paper, we describe the use of the confidential microdata in the LBD to create a series of public use, fully accessible products describing the dynamics of firms and establishments in High Tech industries. This effort is consistent with the goals of the Census Bureau to produce up-to-date economic and social measures to advance informed decision-making in business and society through the development of public use tabulations while also protecting the confidentiality of the underlying data (Jarmin, Louis, and Miranda, 2014).

The existing BDS tabulations, built upon the LBD, contain information on the number of firms and establishments, firm openings and closings, and job creation and destruction by relevant firm characteristics including firm age, size, state, MSA, and sector. However, identifying High Tech businesses in the LBD is not a trivial exercise. In this paper, we review the existing methodologies used to identify High Tech industries, paying special attention to their application and use in BDS data products. These methodologies rely on the identification of inputs and outputs of the innovation process across detailed industries such as R&D investments, proportion of STEM workers, or the technological content of the products. Currently, the BDS statistics are updated annually.[2] A critical element in the evaluation of these methodologies is the ease with which they can be updated and replicated. Relevant factors in this regard include the timelines and coverage of the underlying data that the High Tech classification methodology employs. Some classification methods would require third party data over which the Census Bureau has little control. An equally important consideration when choosing a methodology for identifying High Tech industries is that it produce stable indicators. A methodology that identifies a volatile set of industries is of little use. Finally, an important requirement is that the methodology be flexible enough to identify High Tech industries across the entire economy since it is increasingly the case

[2]Workers in occupations related to Science, Technology, Engineering, and Math (STEM). For examples of such occupations see the Bureau of Labor Statistic's Standard Occupation Classification system Detailed 2010 SOC occupations included in STEM (http://www.bls.gov/soc/ accessed 10/18/2016).

that industries outside of manufacturing can reasonably be considered to be High Tech.

We find considerable differences in the set of High Tech industries identified by different methodologies. Output-based and R&D-based methods almost exclusively identify manufacturing industries due to data limitations. Ad hoc extensions to the methodology attempt to compensate for this inherent limitation but then yield a very broad listing of industries outside of manufacturing so that their value is questionable. Moreover, these output-based methods are relatively difficult to replicate and maintain. Input-based methods, on the other hand, are relatively easy to maintain and replicate given existing and ongoing survey collections. Amongst these methodologies, the one based on occupation data (what workers do) is the most promising. The availability of occupation data across all industries makes it possible to identify a list of High Tech industries that spans all sectors, including information and services industries, using a consistent framework. Despite the significant difference between the methods, there is nevertheless significant overlap in the set of High Tech industries that emerge, particularly in the manufacturing sector. Comparing measures of business dynamism derived from the alternative High Tech definitions shows similar overall patterns in the manufacturing component of High Tech. Consistent with previous findings in the literature, we find overall declines in business dynamism in the High Tech sector after 2000. We find High Tech intensity is directly related to the magnitude of the decline. More intensely-High Tech industries exhibit declines in dynamism that are 20% greater than those experienced by the less intensely-High Tech industries.

Based on concerns about coverage, stability, interpretation, and replicability, we define High Tech industries based on the union of industries with the highest proportion of STEM employment in 2005, 2012, and 2014.[3] Using this classification, we compute the Business Dynamics Statistics of High Tech industries (BDS-HT). These data capture flows of firms, establishments, and employment for High Tech and non-High Tech industries. High Tech industries make up a relatively small portion of the economy–about 4% of firms and 6%

[3]We follow Hecker (2005) and define three levels of intensity. The highest level employs at least 5 times as many STEM workers as the economy average.

of employment in 2014. One of the most pronounced features of the BDS-HT data is the significant increase in entry and young firm activity in the 1990s, which then reverses sharply after 2000.[4]

The rest of the paper proceeds as follows. Section 2 describes different methods for identifying High Tech industries. We evaluate their reproducibility, ease of maintenance and overall strengths and weaknesses. Section 3 compares the list of High Tech industries that emerge across methodologies. We compare selected measures of business dynamism economy wide and for the manufacturing sector. Section 4 presents results from a set of BDS High Tech tables. Section 5 concludes.

2 Identifying High Tech Industries

An interagency workshop held by U.S. statistical agencies in 2004 identified a set of important factors that contribute to the concept of High Tech. These include disproportionately high employment of STEM workers, disproportionately high employment of R&D workers and capital, the production of High Tech products, and the use of High Tech production methods, including the use of High Tech capital goods and services.[5] Efforts to take this definition to the data attempt to classify economic activity based on either the use of High Tech inputs or the production of High Tech outputs. Each approach has advantages and disadvantages with respect to their use in measuring the dynamics of High Tech industries. These differences are primarily related to the availability of data and reliance on arbitrary criteria and thresholds.[6] Table 1 summarizes the different methods of classification and the literature that develops them. In the following sections we detail each method as well as their pros and cons when used in the development of BDS High Tech data.

One conceptual concern to keep in mind before proceeding is that the choice of classification method, whether based on inputs or outputs, will affect the type of activity measured.

[4]This is first documented in Decker, Haltiwanger, Jarmin, and Miranda (2016b).
[5]See Hecker (2005) for a review.
[6]See Kask and Sieber (2002) for insights.

For the purposes of the BDS data products, we will evaluate both input and output methods as well as the amount of arbitrariness involved in each. This is relevant in as much as the statistics we generate might differ across methodologies.

2.1 Input-Based Methodologies

2.1.1 STEM Employment

Hecker (2005) identifies a list of High Tech four-digit 2002 NAICS industries by measuring employment in "technology oriented occupations". This approach is based on the idea that firms that innovate and are engaged in R&D also tend to employ a relatively large number of scientists, engineers, and technicians.[7] In this regard, STEM workers are an input into the innovation process and can therefore be used as a proxy for identifying innovative activity in the economy. Hecker (2005) uses the 2002 Bureau of Labor Statistics (BLS) National Employment Matrix derived from the Occupational Employment Statistics (OES) survey, which provides occupational employment by four-digit NAICS industries, to calculate the proportion of technology oriented workers in each industry. In order for workers in an occupation to be considered technology oriented they must be engaged in R&D, the development of scientific knowledge, and/or the use of that knowledge to develop products and production processes. This set of occupations includes those that apply technology in the design of equipment, processes, and structures, applications related to computers and computation, sales, purchasing, and marketing, as well as quality management and management (Hecker, 2005). Table 2 provides some examples of technology oriented occupations in computer and math sciences, engineering, and physical and life sciences.

Operationalizing this approach requires determining cutoffs for what ought to be considered a "large" number of STEM workers. Hecker (2005) defines High Tech industries as those whose proportion of STEM workers are at least twice the average proportion of

[7] The Office of Technology Assessment's "Technology, Innovation, and Regional Economic Development" report from Sept. 9, 1982 and the NSF special report "Science and Technology Resources in U.S. Industry" NSF 88321, from December 1988 both make this point.

STEM employment across all industries.[8] A total of 46 industries meet this definition, with STEM employment at least twice the average of 4.9 percent in 2002 (see Appendix A for a listing). These industries accounted for 11% of nonfarm wage and salary jobs (14.4 million) in 2002. In 44 of the 46 industries median 2004 earnings were greater than overall median earnings. Similarly, when focusing on occupations, the median earnings in all of the technology oriented occupations was greater than the median for all workers except forest and conservation technicians. The earnings data suggests that the occupations identified as technology oriented indeed capture high skilled employment.

One of the challenges in using input based methods to identify High Tech industries is that they run the risk of biasing the set of industries selected as a function of the type of inputs examined. One of the strengths of the Hecker (2005) methodology is that it relies on labor inputs that are (at least for now) a standard input to the innovation production function and therefore produce results that are less likely to be biased or constrained to certain types of industries.[9] The Service and Information sectors are well represented among the Hecker (2005) High Tech industries–22 of the 46 industries identified as High Tech fall in these sectors, which include computer design services, architectural and engineering services, and data processing and hosting services, amongst others. By focusing on the concentration of STEM employment, the resulting set of High Tech industries is not limited to those industries with relatively high officially-measured R&D expenditures or to those that produce High Tech outputs. In contrast, the STEM method might miss industries that do not rely on technology oriented workers to produce High Tech outputs or industries that utilize other types of High Tech inputs but not technology oriented workers.

One final concern about the STEM methodology is that its reliance on the relative concentration of STEM labor within narrow industries makes it sensitive to industry definitions

[8] He provides further refinement based on the intensity of High Tech. We make use of this refinement to identify industries with the highest proportions of STEM workers; those with at least five times the average.

[9] A mounting concern with innovation metrics is that the increasingly fail to identify certain types of innovations. For example, innovations that have a zero price or do not rely on patents or measured R&D. What used to be the relatively simple task of tracking domestic R&D spending by a small number of U.S. manufacturers has evolved into the need for more broad measures of innovation, science, and technology.

and the boundaries of establishment activity.[10] This is less of a concern in the OES data since establishments, to which 6-digit NAICS codes are assigned, that perform two or more distinct economic activities are treated as two or more separate establishments if separate payroll records are available.[11]

Hecker (2005) groups the 46 identified industries into three High Tech intensity levels. Level I includes 14 industries with a proportion of STEM workers at least 5 times the average. Level II includes 12 industries where the STEM proportion was at least 3 and less than 5 times the average. Level 3 includes 20 industries where the STEM proportion was at least 2 and less than 3 times the average. Between the three levels, all of computer and electronic products industries (NAICS 334) are covered and most of the telecommunications industries within the information sector, four machinery-manufacturing industries, and all but one industry in chemical manufacturing. Moreover, Hecker finds all Level I industries to be R&D intensive and most have some "advanced technology products" as defined by the Census Bureau (NSF, 2002; Census, 2015).

Replicability and Stability

Having the ability to replicate and update the list of High Tech industries is a critical requirement of creating and maintaining a set of statistical products that track the dynamics of High Tech industries.[12] One advantage of the Hecker (2005) methodology is that it can be replicated using more recent industry-occupation employment data from the BLS. The OES industry-occupation employment matrix is available annually between 1997 and 2014. We use the 2012 and 2014 OES data to recalculate STEM concentration ratios for four-digit NAICS

[10]Establishments can engage in multiple types of activities; however, the Census Bureau identifies only the dominant activity, which creates a potential source of noise. Similarly, the Census Bureau identifies headquarter establishments and those offering support services independently from the industry they serve. These types of issues are discussed in greater detail in Census Bureau memorandums related to the classification of auxiliary establishments (http://www.census.gov/eos/www/naics/history/docs/cm_3.pdf accessed 10/18/2016).

[11]For additional information on the sample frame and survey methods see http://www.bls.gov/oes/current/methods_statement.pdf accessed 10/18/2016).

[12]The BLS recently released new statistics for High Tech industries based on the Hecker (2005) methodology. These statistics, however, do not delineate between the different levels or intensities of High Tech, which as we discuss are crucial for producing a stable classification. (http://www.bls.gov/opub/btn/volume-5/the-high-tech-industry-what-is-it-and-why-it-matters-to-our-economic-future.htm accessed 6/20/2016)

industries.[13] This exercise allows us to not only produce measures that more closely resemble current High Tech industries, but also to assess how stable the Hecker (2005) classification methodology is over time. A common concern across classification methodologies is that industries that have been High Tech in the past are no longer considered High Tech and vice versa. Similarly, if there are significant movements across High Tech intensity levels this might inform our choice of the level of aggregation used in the final tabulations. Table 3, Table 4, and Table 5 summarize the results from that analysis.

Table 3 lists the number of 6-digit 2007 NAICS industries identified as High Tech using different vintages of the OES data.[14] These industry counts, broken out by High Tech level, suggest that the total number of industries identified as High Tech is remarkably stable over time. In 2005, there are 173 High Tech industries spanning levels I through III, which dipped to 167 in 2012 then rebounded back to 171 in 2014.[15] Table 3 is useful in determining the changes in the number of industries classified as High Tech within each data vintage and level; however, it is not informative regarding changes in the actual industries covered across years and High Tech intensity levels. Table 4 and Table 5 provide this information. Table 4 shows the number of High Tech industries in the intersection of all three data vintages, as well as the number of High Tech industries specific to each data vintage. We find there are 147 industries common to all three years, which again indicates a high level of stability. We also find 24 industries that were High Tech in 2005 but are no longer High Tech in 2012 and 2014. Similarly, 8 industries are High Tech only according to either the 2012 or 2014 STEM concentration ratios.[16] Table 5 provides the same information by intensity level. Several

[13]It is important to note that the public industry-occupation tables include disclosure suppressions. These suppressions are unlikely to affect our results since the cells being suppressed are often relatively small and are therefore unlikely to impact the ratio of STEM employment within an industry. Moreover, in a comparison of our results to similar efforts performed by BLS that aggregate from undisclosed versions of the data, we find the same sets of industries being identified as High Tech.

[14]The 2012 and 2014 OES industry-occupation employment data use 2012 NAICS industries and Hecker (2005) uses 2002 NAICS industries. We convert these to 2007 NAICS industries using public Census NAICS concordances.

[15]It is unclear whether changes in the number of industries reflects growth or contraction in the number of STEM workers or fragmentation of production activities across establishments and firms. This question is beyond the scope of this paper and we do not address it here.

[16]See Appendix F for list of industries that are classified as High Tech in either 2012 or 2014 but not

things are worth noting. First, out of the 147 industries consistently identified across all three vintages only 107 are classified in the same level of technological intensity in all three years. A significant proportion of the industries classified as High Tech in all three years, 27%, switch technology level over time. There is a considerable amount of fluidity not only in the types of industries that are classified as High Tech but also in the intensity level that they fall in. Second, the more consistently defined level is the highest technological group (level I) as evidenced by the number of industries in common across years versus the industries that are not, 48/11 for Level I vs 27/51 for Level III. Across the three vintages there are 60 unique 6-digit Level I industries, meaning that 80% (48) are consistently Level I over the time period analyzed.[17] Third, a large part of the fluidity across definitions is caused by newly classified Level III industries. All 8 of the industries that only appear in 2012 or 2014 are classified as Level III and none of the 24 industries that only appear in Hecker (2005) are Level I. Because of the STEM methodology's reliance on the mean concentration of technology oriented workers, the Level III group is the most sensitive to small shifts in the distribution of STEM workers across industries. These results indicate that the highest intensity industries (level I) do not appear susceptible to changes in the composition of technology oriented occupations. This is in part driven by the relative concentration of STEM employment in Level I industries. In 2014, Level I industries accounted for 44% of total STEM employment and all three levels accounted for 63% of STEM employment.

To summarize, there are several advantages to using the Hecker (2005) methodology within the BDS High Tech data. First, the Hecker (2005) method relies on readily available data and can be replicated with OES industry-occupation employment data. Second, the methodology produces a stable set of industries over time, especially for the set of industries with a relatively high concentration of STEM workers(Level I). Third, the Hecker (2005) methodology identifies technologies across a wide set of sectors including manufac-

found in the original Hecker (2005) classification.

[17]Note that Table 5 excludes industries that are classified within a given level in some combination of two years. For example, industries that appear as Level I in 2012 and 2014 but not 2005).

turing, services, and information. In this regard, it is important to note that it has already been used to measure the business dynamics of High Tech industries. Hathaway (2013) and Haltiwanger, Hathaway, and Miranda (2014) use the STEM methodology to analyze business dynamics in Level I High Tech industries. Those analyses speak to the appeal of the methodology and provide a successful test case of applying this methodology to the BDS data more broadly.

There are also some disadvantages to using the STEM employment concentrations to identify High Tech industries. First, the cutoffs between the different levels of High Tech industries are still arbitrary. Second, the Hecker (2005) methodology requires as input a curated set of occupations identified as "technology oriented occupations". Such a list of occupations necessarily entails subjective assessment and expert review.[18]

2.1.2 R&D Employment and Intensity

In an earlier effort, Hecker (1999) identifies a list of High Tech three-digit SIC industries using employment in both R&D and technology oriented occupations in the OES surveys from 1993 to 1995. In these years the OES asked employers to report the number of workers engaged in R&D activity. Hecker (1999) identified 31 three-digit SIC industries in which the number of R&D workers and technology oriented occupations accounted for a proportion of employment that was at least twice the average for all industries surveyed. He finds these industries had at least 6 R&D and 76 technology oriented workers per 1,000 workers. The list of 1997 NAICS codes identified as High Tech, concorded from the original list of SIC codes by the Office of Technology Policy and the Census Bureau NSF (2003), are shown in Appendix B. The Hecker (1999) methodology uses more information than Hecker (2005) and has been used by the NSF in their Science and Engineering Indicators publication to

[18]The Standard Occupation Classification Policy Committee (SOCPC) formed a working group that included a number of government agencies to develop a standard definition for STEM occupations. The working group's recommendations were submitted to the Office of Management and Budget (OMB) in 2012. For additional information see "Attachment A: Options for defining STEM occupations under the 2010 Standard Occupational Classification (SOC) system" (http://www.bls.gov/soc/Attachment_A_STEM.pdf accessed 10/18/2016).

define High Tech industries. However, it is important to note that the list of identified industries skews heavily towards manufacturing where R&D as an expense category is more easily measured. Furthermore, this methodology relies on R&D employment data no longer collected in the OES survey and more current NSF R&D employment data from the Business R&D and Innovation Survey (BRDIS) has only limited comparability to the original OES measures. The NSF no longer uses this methodology to identify High Tech firms in favor of the OECD methodology that focuses instead on measures of R&D intensity.

The OECD has produced a classification scheme to identify High Tech industries using R&D intensity. That classification scheme has been used by statistical agencies around the world to develop economic statistics describing High Tech industries. One of the primary goals for developing such a classification scheme was the comparability of relevant statistics across countries.[19] The OECD's R&D intensity classification scheme is based upon Hatzichronoglou (1997)'s measures of direct and indirect R&D intensity for 22 ISIC Rev. 2 manufacturing industries. Behind this approach is the idea that R&D is a key input to innovation and that High Tech firms devote a "high" proportion of expenditures to R&D activities. This input based measure forms the basis for the R&D intensive and High Tech statistics in the OECD's Science, Technology, and Industry (STI) Scoreboards.[20] Hatzichronoglou (1997) uses data from R&D surveys conducted in ten OECD countries between 1980 and 1990 to measure R&D intensity. The structure of these surveys in more recent years is guided by the Frascati Manual of proposed standards and practices (OECD, 2002). Hatzichronoglou (1997) measures direct R&D intensity of an industry as R&D expenditures divided by total industry value added or output. Indirect R&D intensity is measured using input-output tables to account for the technology embodied in intermediate and investment goods. Finally, Hatzichronoglou (1997) uses these measures of direct and indirect R&D intensity to group each of the 22 manufacturing industries into high-technology, medium-

[19]The National Science Foundation uses the OECD's R&D intensive industries to characterize trends in High Tech industries in their Science and Engineering Indicators (NSF, 2010, p. 6-8).

[20]The STI Scoreboard is a compendium of comparable cross country statistics focused on science and technology.

high-technology, medium-low-technology, and low technology classes. The original measures were calculated over a long period, using data between 1980 and 1990, which helped ensure that the indicators would remain stable over time.[21] The high-technology manufacturing group includes: aerospace (3845), computers and office machinery (3825), electronics and communications (3832), and pharmaceuticals (3522).

The OECD has published several updates to the Hatzichronoglou (1997) classification scheme. The list of High Tech industries was updated in 2003 to use ISIC Rev. 3 industries and leverage additional survey data for 1991-1999 (OECD, 2003). These industries are shown in Appendix C. The most recent update, incorporated in the 2015 STI Scoreboards reclassifies High Tech manufacturing industries using ISIC Rev. 4. The High Tech manufacturing industries now include Pharmaceuticals (Division 21), Computer, electronic and optical products (Division 26) and Air and spacecraft and related machinery (Group 303).

Recognizing the inherent limitations of using R&D data, the OECD has also classified High Tech services. Owing to the lack of detailed services data across countries, the 1999 STI Scoreboard identified High Tech service industries as all those that fell in the broad group "Finance, insurance, real estate and business services" (ISIC Rev. 2, division 8). This was updated in the 2001 STI using limited R&D intensity data for services sectors to include: division 64 post and telecommunications, divisions 65-67 finance and insurance, and divisions 71-74 business activities (not including real estate), education, and health (OECD, 2007, p. 124).[22] The most recent revision of the OECD's R&D intensive industry classification was published in 2015 and includes agriculture, mining, utilities, construction, and a broad range of service industries (OECD, 2015a). Given data limitations, the selected set of service industries is relatively broad.

Replicability

[21] See Hatzichronoglou (1997), pg 5.

[22] The NSF Science and Engineering Indicators 2016 describes knowledge intensive service industries as "those that incorporate science, engineering, and technology into their services or the delivery of their services, consisting of business, information, education, financial, and health services." Commercial knowledge intensive services are business, information, and financial services. See remarks to summary page, "ISIC Rev. 3 Technology Intensity Definition", http://www.oecd.org/sti/ind/48350231.pdf.

The OECD recognizes that a different set of industries might result by applying the same methodology to country specific R&D surveys.[23] A U.S. specific list of High Tech industries could be produced using R&D measures found in the NSF's BRDIS data, but there are currently no efforts underway to generate such a list.[24] However, replicating the OECD methodology using the U.S. R&D survey data requires resolving a number of methodological questions. First, R&D surveys are currently skewed away from small and young R&D performers, a disproportionally innovative and dynamic set of firms (Acemoglu, Akcigit, Bloom, and Kerr, 2013). Moreover, there is no clear methodology for grouping the industries by level of R&D intensity. For instance, the R&D intensive manufacturing industries appear to have been grouped into high, medium-high, medium, and low-tech manufacturing industries somewhat arbitrarily.

The NSF's motivation for using the OECD's R&D intensity classification scheme is one of its primary advantages; namely, its international comparability. This is also one of its biggest drawbacks. Due to data limitations international comparability requires identifying a set of ISIC industries that are relatively broad. Moreover, industries listed as R&D intensive in one country may not be R&D intensive in another country (OECD, 2015b, p. 28). As discussed in section 3.1, this is particularly problematic in the U.S. where mapping the High Tech ISIC industries into 6-digit NAICS industries yields many service industries that are not High Tech. Furthermore, as noted in Hatzichronoglou (1997), R&D intensity measures are biased against industries that experience periods of high demand growth in which R&D expenditures increase less rapidly. Finally, the OECD, referencing a wide body of literature indicating that innovation is much broader than R&D intensity, cautioned against interpreting R&D intensive industries as technology-intensive (OECD, 2015a, p. 1).

[23]See remarks to summary page, "ISIC Rev. 3 Technology Intensity Definition", http://www.oecd.org/sti/ind/48350231.pdf.

[24]See NSF Science and Engineering Indicators 2016, pg 44.

2.2 Output-Based Methodologies

2.2.1 Advanced Technologies

The Census Bureau has developed a list of "advanced technology products" (ATP) using the Harmonized System (HS) product codes available in administrative trade data. The goal of this exercise is to produce trade statistics for advanced technology goods. The classification scheme covers both imports and exports and goes back at least to 1989. This set of products includes 10 broad technology areas: biotechnology, life science technologies, opto-electronics, information and communications, electronics, flexible manufacturing, advanced materials, aerospace, weapons, and nuclear technology. In order to be included in one of these categories a product must contain a "significant amount of one of the leading-edge technologies, and the technology must account for a significant portion of the product's value" (NSF, 2002). The Census Bureau maintains concordance files mapping 10-digit HS codes to each technology area.[25] The 2012 crosswalk includes 389 HS codes for exports and 514 HS codes for imports. The BLS assigned the high-level product groups to four-digit 2002 NAICS codes (Hecker, 2005). These industries are shown in Appendix D.

2.2.2 R&D Intensive Products

Hatzichronoglou (1997), in addition to the R&D intensity measures described in Section 2.1.2, also developed a High Tech classification methodology based on products. The methodology used calculations of R&D intensity by groups of products (R&D expenditure/total sales). These SITC Rev. 3 product codes are shown in Appendix E. Relatively little information is available to describe the methods used to define High Tech products. Most OECD publications now rely on industry approaches based on R&D intensity described in the previous section rather than the classification of R&D intensive products.

Limitations

[25] Census Bureau Foreign Trade Statistics, https://www.census.gov/foreign-trade/reference/codes/atp/index.html (accessed 11/24/2015)

Regardless of the source or structure of the product codes identified as High Tech, for the purposes of the BDS statistics, these products would need to be translated to NAICS industries. The BLS translation of the Census ATPs relies on manual curation and does not provide an appropriate level of industry detail for the purposes of the BDS High Tech data. An alternative method of translating the Census ATPs to industry codes would be to leverage the Longitudinal Foreign Trade Transactions Data (LFFTD) to identify the industries in which firms are importing and exporting ATPs. However, mapping product codes associated with trade transactions to specific industry codes is a non-trivial exercise, especially given that large trading firms typically operate across multiple industries. A clear limitation of this approach is that, due to the nature of the trade transaction data, the resulting list would be heavily skewed towards manufacturing industries.

3 Comparing Methods

3.1 Industry Coverage

Despite the very different methodologies used to create the STEM concentration, R&D intensity, and product classification schemes, there is significant overlap in the resulting set of industries identified as High Tech, particularly within manufacturing. For example, all four ISIC Rev. 4 High Tech industries used in the 2015 OECD STI report, pharmaceuticals, computer, electronic and optical products, and air and spacecraft and related machinery, are represented in the Hecker (2005) Level I industries as well as the recalculated STEM 2012 and STEM 2014 Level I industries. Moreover, industries identified as High Tech in earlier OECD STI reports such as medical, precision and optical instruments, and radio, TV, and communications equipment are also captured in the Level I STEM concentration industries. Similarly, 9 of the 16 four-digit 2002 NAICS industries covered under the Census Bureau's "advanced technology products" appear in the Hecker (2005) Level I industries. Three of the remaining seven industries are found in Level II and none in Level III.

In order to perform a more thorough comparison of the makeup of the High Tech industries identified by different methodologies we translate each set from its native industry scheme to 6-digit 2007 NAICS codes. As with any application requiring the translation across industrial classification systems, there is some measure of ambiguity in the final concordance. Specifically, the Hecker (1999) High Tech industries include a combination of 4, 5, and 6-digit 1997 NAICS industries. We translate the 4 and 5-digit industries to the 6-digit 1997 NAICS industry classification system, then concord at the 6-digit level to the 2007 NAICS industry classification system using Census Bureau NAICS concordances.[26] The OECD (2003) High Tech industries include a combination of 2, 3, and four-digit ISIC Rev. 3 industries, which we first translate to the four-digit ISIC Rev. 4. industry classification system then to the 6-digit 2007 NAICS industry system using the United Nations concordances.[27] The recently revised OECD (2015a) R&D intensity industry classification includes 2 and 3-digit ISIC Rev. 4 industries, which we translate to 6-digit 2007 NAICS industries. At each step, it is possible that one industry in the original classification maps to multiple industries in the new classification and vice versa. For example, in the case of the OECD's R&D intensive classifications, some of the original ISIC industries are relatively aggregate, such as 31 Medical, precision, and optical instruments. This high level of aggregation means one ISIC industry will ultimately map to many 6-digit 2007 NAICS industries. In addition, due to the many-to-many nature of this translation, it is not always possible to assign a unique ISIC based High Tech intensity level to a given NAICS (i.e. the same NAICS code can be classified as high and medium-high according to its mapping to different ISIC codes). In such cases, we select the lowest classification level.[28]

Table 6 shows the number of 6-digit industries identified by each methodology as High

[26] See http://www.census.gov/eos/www/naics/concordances/concordances.html (accessed 01/21/2016)

[27] http://unstats.un.org/unsd/cr/registry/regot.asp?Lg=1 (accessed 01/21/2016)

[28] As an example, in the original ISIC Rev. 3 codes, ISIC 33 is high tech, ISIC 29 is medium-high tech, and ISIC 36 is low tech. Once mapped to four-digit ISIC Rev. 3.1 and concorded to the 6-digit 2002 NAICS industries, all three 2-digit ISIC Rev. 3 industries map to 2002 NAICS 238290 Other Building Equipment Manufacturers, meaning that industry is simultaneously classified as high, medium-high, and low tech. In this case, we select the lowest technology group, meaning 238290 would be considered a low technology industry.

Tech across major sectors. All methodologies identify a large number of manufacturing industries as High Tech. The Census ATP classification focuses heavily on manufacturing. The STEM concentration methods and OECD definitions also identify a significant number of information and services industries as High Tech. The 2003 vintage of the OECD method identifies 50 industries within the Finance, Insurance, and Real Estate, far more than any other method. Additionally, the 2003 OECD classification identifies twice as many industries as High Tech compared to the STEM methods. Note that the services sector here includes both professional services (54) and other non-public administration services (81). Examining the detailed industries included by the 2003 OECD classification we begin to see the ambiguity and noise inherent in translating 2-digit ISIC Rev. 3 codes to 6-digit 2007 NAICS codes. For example, within the services sector, the OECD method identifies six industries under 812 Personal and Laundry Services as well as Grantmaking Foundations (813211) as High Tech.[29] Also, under the Arts & Accommodation sector, the OECD method includes Agents and Managers for Artists, Athletes, Entertainers, and Other Public Figures (711410) as High Tech. The 2015 vintage of the OECD classification has backed away from the broad identification of services as High Tech and now identifies many fewer service industries. It is now very much a manufacturing sector classification. Next, we use the alternative classifications to build selected business dynamics statistics and explore their broad patterns. Before doing so we turn to a discussion of issues related to the development and interpretation of BDS High Tech data.

3.2 Methodological Considerations

One of the most compelling dimensions of the BDS data is the length of its time series. The BDS data are available back to 1977. This allows researchers to examine not only how the U.S. economy is changing over time but also how it responds to specific shocks.

[29] Coming through ISIC Rev 3 2-digit 74 to four-digits 7412, 7413, 7414, 7421, 7422, 7430, 7491, 7492, 7493, 7494, 7495, 7499 which concords to ISIC Rev 3.1 9309, 7494, 5260, 6420, which concords to 2007 NAICS 812.

However, producing BDS High Tech data over this long time period introduces a number of issues related to interpretation and longitudinal consistency of the series. All of the methods described in the previous sections identify a set of High Tech industries at a point in time (or short window of time). The STEM methodology relies on point in time estimates of occupational concentrations across industries and the OECD R&D intensity measures relies on point in time estimates of R&D expenditures and industry output. Moreover, though the classifications produced by the STEM methodology are relatively stable, the definition of High Tech does change over time. Industries that were High Tech a few years ago may no longer be High Tech today while new High Tech industries are appearing that did not exist before. This problem only becomes more pronounced across several decades. Due to these changes over time, we might be concerned that the BDS High Tech measures will be sensitive to the choice of vintage. Put differently, which vintage is used has implications for the interpretation of the dynamism measures we create since they will be specific to a particular set of industries and therefore a specific set of technologies existing at a point in time.[30] There is a trade off, therefore, between how often to update the classification scheme and how broad to make the set of industries we track. Broader measures will be more stable over time but noisier and less descriptive of High Tech sectors. Alternatively, one might focus on a set of industries that is relative stable (e.g. Level I) and pool the industries classified as High Tech Level I across several years. Ultimately, how much we need to worry about this depends on how sensitive the business dynamics statistics are to the vintage used, which we turn to now.

Figure 1 shows firm entry rates in High Tech manufacturing according to several different aggregations of the STEM classifications using 2005, 2012, and 2014 occupation data. We focus first on the Manufacturing sector since this is the one sector where all the alternative measures are most directly comparable. Firm entry rates are produced for Level I industries (the most STEM intense industries) as well as all three High Tech levels combined. Turning

[30]Note that this is less of a concern for the STEM Level I industries, which are more stable over time.

to Level I industries, several patterns are worth highlighting. First, all series, regardless of vintage, show considerable decline in entry rates over time, which is consistent with well-established economy wide patterns. Second, there is a slight difference across the different STEM groupings in the 1980s but by 1990 the rates converge. This is to be expected given that as we move back in time the more current classifications become less representative of High Tech sectors in the past. The more relevant takeaway, however, is that the Level I STEM measures are not very sensitive to the choice of vintage. The difference between the STEM measures is less than 1 point at its widest point. This is not surprising given the significant overlap in industry coverage across STEM methodologies within the Level I group. Recall 80% of Level I industries appear in 2005, 2012, and 2014 STEM High Tech industries.[31] The 2012 and 2014 Level I measures are actually identical because, for the manufacturing sector, they happen to include the same 6-digit NAICS industries.

The patterns for the broader technology classification, which combine all High Tech levels, are similar. We find small differences across vintages with the largest gap being less than 1 entry point. Importantly, comparison of the narrower versus the broader High Tech definitions shows Level I High Tech manufacturing industries experience a considerably steeper decline in firm entry rates. The decline in Level I High Tech industries is about 20% steeper than the combined groupings. The decline for Level I industries is roughly 9 entry points versus roughly 7.5 entry points for the combined set of Level I, Level II, and Level III industries.

Figure 2 shows firm entry rate statistics for all High Tech industries including those outside of manufacturing. The patterns are quite different from those in manufacturing. High Tech industries outside of manufacturing experienced a very dynamic period of high entry rates in the mid-1990s that was reversed in the late 1990s and early 2000s.[32]. We also see that Level I High Tech industries are particularly sensitive to the underlying economic forces

[31] Other measures of dynamism across other sectors show similar patterns of agreement among the STEM and R&D methodologies.

[32] This is in agreement with findings in (Decker, Haltiwanger, Jarmin, and Miranda, 2016b)

driving these changes since they experience both steeper growth and decline. This suggests that using a broader classification scheme might prevent us from recognizing shocks affecting the more intensely High Tech industries. Overall, across all of the STEM classifications, the time series patterns show a high correlation rate. On average across the series the correlation is over 0.9 for both manufacturing and economy wide.

Figure 3 compares High Tech manufacturing firm entry under both the OECD 2003 and OECD 2015 methodologies as well as the Census Bureau Advanced Technology Products classification. As with STEM manufacturing, the OECD manufacturing High Tech classifications show a steep trend decline in entry rates over this period. The vintage used for defining the OECD High Tech industries, at least for manufacturing, does not matter much despite more than a decade between the data underlying the two classifications. It is not surprising that these series track each other given the level of overlap between the OECD industries, ATP industries, and STEM industries.

Figure 4 shows firm entry rate statistics for all High Tech industries, including those outside of manufacturing, based on the OECD and Census classifications. As before, we see the OECD classification misses some important dynamics captured by the STEM classification. The average correlation between the OECD and Census ATP series is again over 0.9 for manufacturing and about 0.8 for all High Tech industries Figure 1 and Figure 2 suggest that, in the case of the STEM concentration methodologies, the choice of vintage does not have as great an impact on the measures of dynamism as one might have expected. The more relevant dimension is the breath of industries covered by the different methods. Despite the stability of these measures, the choice of vintage still matters for their interpretation. The more recent vintages are not as descriptive of the High Tech industries of the past. The Hecker 2005 and OECD 2003 classifications more accurately reflect High Tech activity in the 1980s and 1990s relative to the more recent STEM 2014 and OECD 2015 classifications. The older vintages are also not as representative of newer technologies and industries. An important consideration then is how frequently the classification scheme could or should

be updated. The BLS industry-occupation employment matrix is published annually and updates to the OECD R&D intensity measures occur every few years. Depending on the methodology selected, a new classification scheme could be generated either annually or every few years. With each set of new industries, a new time series would be produced. However, updating the High Tech classification frequently creates potential disclosure problems.[33] Solutions to this problem include creating noise infused statistics, lower frequency updates, or using differential privacy methods.[34]

4 BDS High Tech Statistics

After weighing issues related to the coverage, stability, interpretation, and replicability of different classification methodologies, we chose the union of Level I industries according to the concentration of STEM employment in 2005, 2012, and 2014 as our preferred methodology for the BDS High Tech data. One motivating factor for this decision is that the STEM concentration method captures High Tech activities outside of the manufacturing sector. As shown in Section 2.1.1, the Level I industries change very little over time. Moreover, pooling High Tech industries across years will make the classification scheme even more stable. This pooling also yields a relatively clear interpretation of the final data series. The resulting statistics capture the current and historical performance of industries that were High Tech in the past decade. They do not capture the business dynamics of industries that were High Tech in the late 1970s and 1980s but would no longer be considered High Tech. In order to avoid issues of complimentary disclosure and discontinuities in definitions, we intend to fix the classification scheme for several years as opposed to updating annually as new industry-occupation data becomes available.

[33]The differential industries would be subject to disclosure analysis, greatly increasing the burden on analysts and disclosure officials and possibly the number of suppressed cells.

[34]Noise infusion is a disclosure protection method that obviates the need for secondary cell suppression by ensuring that no single cell is disclosive (Evans, Zayatz, and Slanta, 1996). Differential privacy encompasses methods that provide a robust, meaningful, and mathematically rigorous definition of privacy, together with a computationally rich class of algorithms that satisfy those definitions (Dwork, Roth, et al., 2014).

The STEM union includes 15 four-digit 2007 NAICS industries shown in Table 7. The STEM union covers industries in mining, manufacturing, information, and professional services. Comparing this list of industries to those found in Hecker (2005), we see two additional industries, Oil and Gas Extraction (2111) and Other Information Services (5191). Though this set of industries is not drastically different from those in Hecker (2005), updating the list based on more recent occupation data allows the list to be more broadly representative of High Tech industries in the 2005 to 2014 time period. This classification yields a set of industries that play a particularly important role in reallocation by introducing new products and services. The NSF published figures from the 2011 BRDIS that measure the share of businesses that introduce new goods or services between 2009-2011. These include estimates for 9 of the 15 four-digit NAICS industries we classify as High Tech. On average, 40% of the businesses in these industries introduced a new product or service compared to 9% for all industries and 22% for manufacturing.[35]

In addition to identifying industries of interest, we must also deal with issues related to changing industrial classification systems across years in the underlying LBD microdata. Prior to 2002, establishments in the LBD were classified by SIC. Between 2002 and 2014 establishments are assigned NAICS codes using some combination of the 1997, 2002, and 2007 vintages depending on the year in question. For the construction of the BDS High Tech data, we leverage the work by Fort and Klimek (2016) to apply consistent industrial classification codes to all establishments in the underlying microdata. Their methodology uses information from Economic Censuses, administrative data, and a variety of industry concordances to do this assignment. Ultimately, they are able to assign a single complete 6-digit 2007 NAICS code to all establishments in the data for years 1976 through 2014. While Fort and Klimek (2016) allow industry classifications for each establishment to vary over time, for our purposes we select a single, time-invariant industry code for each establishment.[36]

[35]See https://www.nsf.gov/statistics/2016/nsf16308/overview.htm for details.

[36]Industry codes can change for a variety of reasons some legitimate (e.g. changes to production lines) some less so (e.g. errors and non response).

To do this, we select either the industry code associated with the most recent Economic Census year or, in cases where the establishment was never covered in an Economic Census, the modal industry code.[37] Having these consistent industry codes allows us to use our list of High Tech industries to produce business dynamics statistics for all years covered in the BDS.

The BDS High Tech data, as with the core BDS tables, measure job creation and destruction from establishment entry and exit as well as expansion and contraction across relevant characteristics including firm age and firm size. Job creation and destruction, JC and JD, are defined following Davis, Haltiwanger, Schuh, et al. (1998) as follows:

$$JC_{st} = \sum_{i \in s, g_{it} \geq 0} (E_{it} - E_{it-1}) \tag{1}$$

$$JD_{st} = \sum_{i \in s, g_{it} \leq 0} |(E_{it} - E_{it-1})| \tag{2}$$

For employment E of establishment i in year t for each group s. Establishment-level growth g_{it} is defined as follows.

$$g_{it} = \frac{(E_{it} - E_{it-1})}{X_i} \tag{3}$$

Where the establishment-level employment denominator, X_{it}, is defined as follows.

$$X_{it} = 0.5 * (E_{it} + E_{it-1}) \tag{4}$$

The net change in employment, NET, for establishments in group s then satisfies the following identity.

$$NET_{st} = \sum_{i \in s} (E_{it} - E_{it-1}) = JC_{st} - JD_{st} \tag{5}$$

Aggregate growth rates, JCR, JDR, and $NETR$, can then be calculated using the

[37] This is consistent with the methodology currently implemented in the BDS.

following.

$$JCR_{st} = \frac{JC_{st}}{X_{st}} \tag{6}$$

$$JDR_{st} = \frac{JD_{st}}{X_{st}} \tag{7}$$

$$NETR_{st} = \frac{NET_{st}}{X_{st}} = \frac{(JC_{st} - JD_{st})}{X_{st}} \tag{8}$$

Where the aggregate denominator, X_{st}, is defined as follows.

$$X_{st} = \sum_{i \in s} X_{it} \tag{9}$$

These growth rate measures are standard in analysis of establishment and firm dynamics. It shares useful properties of log differences but also accommodates entry and exit (Davis, Haltiwanger, Schuh, et al., 1998; Törnqvist, Vartia, and Vartia, 1985). These measures are equivalent to the ones used in the construction of the existing BDS statistics.[38]

We provide a preview of selected statistics here to demonstrate the types of measures included in the BDS High Tech data. The share of firms in High Tech industries is relatively small but it has grown considerably. On average between 1977 and 2014 High Tech industries account for about 3.5% of firms and 6.2% of employment. Figure 5 shows how the share of firms and employment in High Tech industries has changed over time. Since the classification of High Tech is based upon industry, and industry is an establishment level characteristic, firms may be classified as both High Tech and non-High Tech. With that in mind, High Tech firm share in Figure 5 measures the total number of firms with at least one establishment in a High Tech industry divided by the total number of firms in the economy. Since employment occurs at the establishment level, High Tech employment share simply captures the percent

[38]See http://www.census.gov/ces/dataproducts/bds/methodology.html for details on the BDS statistics.

of employment in High Tech industries. The figure suggests that while the overall share of firms in High Tech industries has more than doubled since 1980, High Tech employment share has been relatively flat and trending slightly up since 2005.

Figure 6 shows the job creation rates in High Tech and non-High Tech industries. We see three distinct periods in the High Tech job creation rate series. First, a declining trend similar to that found in non-High Tech industries from 1977 through the early 1990s, followed by a marked increase in the High Tech job creation rate during the 1990s, and finally a return to decline in the post 2001 period. Much of these trends can be explained by changes in firm entry. Figure 7 shows the firm entry rate for the High Tech and non-High Tech sectors. In the non-High Tech sector, the firm entry rate exhibits the familiar secular decline found in many recent studies (Decker, Haltiwanger, Jarmin, and Miranda, 2014). In High Tech, on the other hand, the firm entry rate undergoes a sharp increase from 1992 to 1998, then returns to the declining trend, actually falling faster than non-High Tech by the end of the series.[39] Rising job creation and entry in High Tech during the 1990s corresponds to a period of rapid technological progress in information technology producing sectors, which dramatically reduced prices. Across the economy, firms increased their investments in information technology capital. These investments accounted for upwards of three-quarters of post-1995 labor productivity growth (Oliner and Sichel, 2003; Jorgenson, Ho, and Stiroh, 2002). The responsiveness of businesses to productivity shocks also increased during this time leading to improvements in allocative efficiency and productivity growth Decker, Haltiwanger, Jarmin, and Miranda (2018). The use of information technology played a transformational role in the 1990s by decreasing costs, improving business processes and work practices, and introducing new products or improving existing products (Brynjolfsson and Hitt, 2000).

Even more dramatic than the High Tech boom was the subsequent bust. In Figure 8 we show High Tech job creation as well as job destruction rates. The fall in the High Tech job creation rate is accompanied by a sharp increase in the High Tech job destruction rate in

[39]These patterns have been documented in Haltiwanger, Hathaway, and Miranda (2014); Decker, Haltiwanger, Jarmin, and Miranda (2016a,b).

2002 following the collapse of the dot-com bubble. In analyses of the microdata underlying these tabulations, we find these dynamics are driven in large part by the Computer Systems Design and Related Services (5415) and Architectural, Engineering, and Related Services (5413) industries. These industries provide services such as custom computer programming, computer systems design, computer facilities management and other specialized design services to other firms. The sharp declines in High Tech occurred while there was a fundamental shift in productivity growth. In the post-2000 period the contribution of both information technology production and use to productivity growth began to recede from their extraordinary rates of the 1990s (Jorgenson, Ho, and Stiroh, 2008; Fernald, 2014). Interestingly, the post-2000 declines in these industries, which we identify based on their relatively high levels of STEM employment, also coincides with falling demand for high skilled labor (Beaudry, Green, and Sand, 2016).

To further investigate how these patterns vary across different types of firms we compare job creation by firm age. Figure 9 shows the share of job creation accounted for by startups, young firms, and middle-aged firms. In Figure 9, the increased contribution of startups and young firms to job creation in High Tech in the 1990s is dramatic, rising by nearly 5 and 10 points respectively. The decline is equally dramatic in the years leading up to the Great Recession. Decker, Haltiwanger, Jarmin, and Miranda (2016a) show young High Tech firms displayed a pattern of increased then decreased responsiveness to idiosyncratic shocks during this time, which contributes to the increase and then decline in productivity growth. The contribution to job creation from firms 6 to 10 years old fell through 2002. In the post 2000 period all three series appear to converge.

Though focusing on young firm activity can be informative, the BDS High Tech data allow us to examine these trends across the full range of firm age and size categories. Figure 10 and Figure 11 show the average job creation rates for High Tech industries for the 1983-1992, 1993-2001, and 2001-2014 periods. These periods correspond to the decline, rise, and fall in job creation rates (Figure 6) and share of young firms (Figure 9). In Figure 10 we

see that the rise and fall of job creation was much more dramatic for young firms relative to older firms. Interestingly, job creation for the oldest group of firms (those born before 1977) has grown slowly throughout. Figure 11 shows the job creation rates for High Tech industries by firm size using the same time periods. Here we see that the job creation rate of the smallest High Tech firms, those between 1 and 4 employees, has been in decline over the entire period with the boom period in the late 1990s simply slowing that trend. Medium sized firms experienced patterns similar to young firms–briefly increasing job creation rates in the late 1990s then returning to a declining trend thereafter. The largest High Tech firms, as with the older firms, saw rising job creation rates throughout. These patterns highlight the heterogeneity in the trends in High Tech across different types of firms.

Importantly, these statistics highlight the dramatic changes occurring in the High Tech sector over time and how those changes differ from trends in the non-High Tech sector. Figure 12 shows the firm age distribution for High Tech and non-High Tech sectors and its evolution between 1987 and 2014. High Tech is increasingly dominated by older firms. As the sector matures, the share of young firms (those up to 5 years old) declined from 56% in 1987 to 37% in 2014. The decline in startups and young firm activity follows the secular trends in the non-High Tech sector. The temporary boom during the 1990s is clearly identifiable in the High Tech figure. Figure 13 shows the employment weighted firm shares by firm size groups in the High Tech and non-High Tech sectors. High Tech employment is increasingly concentrated in smaller firms. As the sector matures the share of employment at large firms (those with 5,000 or more employees) declined from 67% in 1977 to 44% in 2014–a 23 point decline.[40] These trends are markedly different from the trends we observe in the rest of the economy. In the non-High Tech sector employment activity is slowly shifting towards the largest firms. The shift away from small firms, 6 points between 1977 and 2014 from 40% to 34%, largely reflects a decline in young firm activity, many of whom tend to be small.

Digging deeper into the nature of the changing nature of the High Tech sector we find

[40]During this time the share of High Tech firms with 5,000 or more employees fell from 0.85% to 0.33%.

dramatic changes in average establishment size. Figure 14 shows the average establishment size for firms in High Tech industries and Figure 15 shows average establishment size for non-High Tech industries by firm size groups. In Figure 14 we see that the average size of establishments of large High Tech firms has experienced a 71% decline from its peak in 1983. These declines are much sharper than the declines observed in the rest of the economy where establishment size at large firms declined by only 17% in the same period. The sharp decline in establishment size at large firms is potentially due to shifts in industry composition in High Tech over this period–with production moving away from manufacturing and towards services and information service providers along with changes in production technology that allow for increased fragmentation of production.

Finally, the BDS High Tech data also allows for the exploration of geographic variation in High Tech firm and employment measures. We might expect employment in the High Tech sector to be relatively concentrated. A number of studies have demonstrated the importance of regional clusters and the geographic concentration of innovative activity (Fujita, Krugman, and Venables, 1999; Delgado, Porter, and Stern, 2014). Figure 16 shows the average share of High Tech employment by state between 2012 and 2014. There is significant geographic variation in the percent of employment in High Tech industries ranging from Virginia with 11.1% to Mississippi with 2.4%. We find the highest proportion of High Tech employment in Virginia, Washington, Maryland, and Massachusetts, followed by California and the District of Columbia. For High Tech firm share, we find the top three states have an average of 6.4% of firms in High Tech industries whereas states at the bottom average 3%. These geographic differences largely persist over time.

To summarize, these figures demonstrate the ways in which the BDS High Tech data can be used to examine the business dynamics in High Tech industries. High Tech industries experienced a surge of entry and young firm activity in the 1990s that was abruptly reversed in the early 2000s. Since then, job creation rates have returned to previous trend declines. The 1990s then appear to represent a break within an otherwise broader secular decline. Firm

entry rates in the High Tech and non-High Tech sectors are now converging. It appears that the boom-bust phenomena in High Tech is concentrated among young relatively small firms–larger older High Tech firms weathered the early 2000s much better than did young small firms.

The BDS High Tech statistics shed light on important structural changes in the U.S. economy. The share of firms in High Tech is increasing over time. However, the share of employment in High Tech industries has remained stable over the last 30 years. The dynamics underlying these trends are dramatically different across firm age and size categories. High Tech is increasingly dominated by older and smaller firms. This is because employment is shifting away from large firms. The average size of establishments for the largest High Tech firms has declined significantly. This is in stark contrast with the non-High Tech sector where there is a significant and well-documented shift in employment away from small firms and establishment size has not declined as much. Finally, geographic variation in High Tech activity is significant and persistent.

5 Conclusion

The reallocation of labor and capital is an important driver of economic growth and rising living standards. Businesses in the High Tech sector play a particularly important role in this reallocation process by introducing new and improved products and services. This paper describes an extension of the existing BDS statistics that improves our understanding of this important and dynamic portion of the economy. Despite significant interest in measuring High Tech economic activity, there is little consensus on how best to identify High Tech businesses. We survey some of the existing methodologies, including input and output based methods, with an eye to their application in producing BDS statistics for High Tech industries. Based on coverage, stability, interpretation, and replicability, we chose the union of Level I industries according to the proportion STEM employment in 2005, 2012, and 2014 as

our preferred methodology for identifying High Tech industries in the BDS data. Finally, we use that definition to describe the firm, establishment, and employment flows in High Tech industries. These new data provide measures of job creation and destruction, entry and exit in High Tech industries by a number of relevant firm characteristics including firm age and firm size. In this paper, we preview some of the resulting statistics and demonstrate their utility in better understanding trends in High Tech. We show the High Tech sector experienced dramatic structural shifts during the last 30 years with a short burst of extraordinary dynamism in the 1990s.

References

Acemoglu, D., U. Akcigit, N. Bloom, and W. R. Kerr (2013) "Innovation, reallocation and growth," Discussion paper, National Bureau of Economic Research.

Beaudry, P., D. A. Green, and B. M. Sand (2016) "The great reversal in the demand for skill and cognitive tasks," *Journal of Labor Economics*, 34(S1), S199–S247.

Brynjolfsson, E., and L. M. Hitt (2000) "Beyond computation: Information technology, organizational transformation and business performance," *Journal of Economic perspectives*, 14(4), 23–48.

Census (2015) "Census Bureau Foreign Trade Statistics," Discussion paper, US Census Bureau.

Davis, S. J., J. C. Haltiwanger, S. Schuh, et al. (1998) "Job creation and destruction," *MIT Press Books*, 1.

Decker, R., J. Haltiwanger, R. Jarmin, and J. Miranda (2014) "The Role of Entrepreneurship in U.S. Job Creation and Economic Dynamism," *The Journal of Economic Perspectives*, 28(3), 3–24.

Decker, R. A., J. Haltiwanger, R. S. Jarmin, and J. Miranda (2016a) "Declining business dynamism: Implications for productivity," *Brookings Insitution Hutchins Center Working Paper No*, 23.

——— (2016b) "Where has all the skewness gone? The decline in high-growth (young) firms in the US," *European Economic Review*, 86, 4–23.

Decker, R. A., J. C. Haltiwanger, R. S. Jarmin, and J. Miranda (2018) "Changing business dynamism and productivity: Shocks vs. responsiveness," Discussion paper, National Bureau of Economic Research.

Delgado, M., M. E. Porter, and S. Stern (2014) "Clusters, convergence, and economic performance," *Research policy*, 43(10), 1785–1799.

Dwork, C., A. Roth, et al. (2014) "The algorithmic foundations of differential privacy," *Foundations and Trends® in Theoretical Computer Science*, 9(3–4), 211–407.

Evans, T., L. Zayatz, and J. Slanta (1996) "Using noise for disclosure limitation of establishment tabular data," in *Proceedings of the Annual Research Conference, US Bureau of the Census, Washington, DC*, vol. 20233, pp. 65–86.

Fernald, J. (2014) "A quarterly, utilization-adjusted series on total factor productivity," Federal Reserve Bank of San Francisco.

Fort, T. C., and S. D. Klimek (2016) "The effect of industry classification changes on us employment composition," *Tuck School at Dartmouth*.

Fujita, M., P. R. Krugman, and A. J. Venables (1999) *The spatial economy: cities, regions and international trade*, vol. 213. Wiley Online Library.

Haltiwanger, J., I. Hathaway, and J. Miranda (2014) "Declining business dynamism in the US high-technology sector," .

Hathaway, I. (2013) "Tech starts: High-technology business formation and job creation in the United States," .

Hatzichronoglou, T. (1997) "Revision of the high-technology sector and product classification," .

Hecker, D. (1999) "High-technology employment: a broader view," *Monthly Lab. Rev.*, 122, 18.

Hecker, D. E. (2005) "High-Technology Employment: a NAICS-Based Update," *Monthly Lab. Rev.*, 128, 57.

Jarmin, R. S., T. A. Louis, and J. Miranda (2014) "Expanding the role of synthetic data at the US Census Bureau," *Statistical Journal of the IAOS*, 30(2), 117–121.

Jarmin, R. S., and J. Miranda (2002) "The longitudinal business database," .

Jorgenson, D. W., M. S. Ho, and K. J. Stiroh (2002) *Projecting productivity growth: lessons from the US growth resurgence*, vol. 20036. Washington, DC.

——— (2008) "A retrospective look at the US productivity growth resurgence," *Journal of Economic Perspectives*, 22(1), 3–24.

Kask, C., and E. Sieber (2002) "Productivitiy Growth in High-Tech Maufacturing Industries," *Monthly Lab. Rev.*, 125, 16.

NSF (2002) "Science and Engineering Indicators," Discussion paper, National Science Foundation.

——— (2003) "Science and Engineering Indicators," Discussion paper, National Science Foundation.

——— (2010) "Science and Engineering Indicators," Discussion paper, National Science Foundation.

OECD (2002) "Frascati Manual: Proposed Standard Practice for Surveys on Research and Experimental Development, 6th edition," Discussion paper, Organisation for Economic Co-operation and Development.

——— (2003) "Science, Technology and Industry Scoreboard 2003, Annex 1," Discussion paper, Organisation for Economic Co-operation and Development.

——— (2007) "Knowledge-Based Industries. Paris: Directorate for Science, Technology, and Industry/Economic Analysis Statistics," Discussion paper, Organisation for Economic Co-operation and Development.

——— (2015a) "The R&D Intensity of Economic Activities in OECD Countries: Proposal for a New Classification for Industry and Services. Directorate for Science, Technology and Innovation; Committee for Scientific and Technology Policy," Discussion paper, Organisation for Economic Co-operation and Development.

——— (2015b) "Science, Technology and Industry Scoreboard 2015," Discussion paper, Organisation for Economic Co-operation and Development.

Oliner, S. D., and D. E. Sichel (2003) "Information technology and productivity: where are we now and where are we going?," in *Technology, Growth, and the Labor Market*, pp. 41–94. Springer.

Schumpeter, J. A. (1942) *Capitalism, socialism and democracy.* New York: Haper & Brothers.

Stinson, M., K. White, and J. Lawrence (2017) "Upcoming Improvements to the Longitudinal Business Database and the Business Dynamics Statistics," Discussion paper, U.S. Census Bureau, Center for Economic Studies.

Stoneman, P., and G. Battisti (2010) "The diffusion of new technology," *Handbook of the Economics of Innovation*, 2, 733–760.

Törnqvist, L., P. Vartia, and Y. O. Vartia (1985) "How should relative changes be measured?," *The American Statistician*, 39(1), 43–46.

Tables

Table 1: Methodologies of Identifying High Tech Industries

Method	Literature
Inputs Based	
STEM and R&D Employment	Hecker (1999), Hecker (2005), Hathaway (2013), Haltiwanger, Hathaway, and Miranda (2014)
R&D Intensity (direct and indirect)	Hatzichronoglou (1997), OECD (2007), NSF (2010)
Output Based	
Production of High Tech Products	NSF (2002), Hatzichronoglou (1997)

Table 2: Example High-Tech Occupations

SOC Code	Occupations
Computer and Math Sciences	
11-3020	Computer and information systems managers
15-0000	Computer and mathematical scientists
Engineering and related	
11-9040	Engineering managers
17-2000	Engineers
17-3000	Drafters, engineering, and mapping technicians
Physical and Life Sciences	
11-9120	Natural sciences managers
19-1000	Life scientists
19-2000	Physical scientists
19-4000	Life, physical, and social science technicians

Source: Hathaway (2013), Bureau of Labor Statistics.
Notes: For full list, see Table 3 in Hecker (2005).

Table 3: Number Industries by Vintage and Group

	Vintage		
Group	Hecker (2005)	STEM 2012	STEM 2014
Level I	55	50	52
Level II	58	54	60
Level III	60	63	59
Total	173	167	171

Source: Hecker (2005), BLS 2012 and 2014 industry-occupation employment matrices.
Notes: All counts are of 6-digit 2007 NAICS industries. Original Hecker (2005) industries were 3 and 4-digit 2007 NAICS industries. These have been converted to 6-digit industries for comparability.

Table 4: Number of Industries Overlapping and Exclusive to each Vintage

Vintage	Industries
All 3 Years	147
Hecker (2005) Only	24
STEM 2012 Only	1
STEM 2014 Only	7
Combination of Two Years (Residual)	19
Total	198

Source: Hecker (2005), BLS 2012 and 2014 industry-occupation employment matrices.
Notes: Includes counts of 6-digit 2007 NAICS industries associated with different combinations of High Tech methodologies. The total captures the count of unique industries across all three vintages. For a detailed list of industries only appearing in 2012 and 2014 see Appendix F. The residual category captures all industries that are observed in some combination of two years2005, 2012, and 2014. We include the residual category for completeness.

Table 5: Number Industries Overlapping and Exclusive by Vintage and Group

	Level I	Level II	Level III
All 3 Years	48	32	27
Hecker (2005) Only	6	7	26
STEM 2012 Only	2	4	11
STEM 2014 Only	3	5	14

Source: Hecker (2005), BLS 2012 and 2014 industry-occupation employment matrices.
Notes: Includes counts of 6-digit 2007 NAICS industries associated with different combinations of High Tech vintages.

Table 6: High Tech Industry Counts by Major Sector and Methodology

Major Sector	Hecker (2005)	STEM 2012	STEM 2014	Census ATP	OECD (2003)	OECD (2015)
Agriculture, Forestry, Fishing and Hunting	2	.	.	.	.	.
Arts & Accommodation	.	.	.	.	1	.
Construction	.	.	.	.	.	.
Education & Health	.	.	.	.	2	.
FIRE	2	8	7	.	50	.
Information	8	11	11	1	14	2
Management & Support	5	3	3	.	27	.
Manufacturing	111	106	105	90	133	236
Mining & Utilities	8	8	15	.	4	.
Retail & Wholesale	7	7	7	.	.	.
Services	26	22	22	7	57	8
Transport & Warehousing	4	2	1	.	5	.
Total	173	167	171	98	293	246

Notes: Counts are of 6-digit 2007 NAICS industries. For Hecker (2005), STEM 2012, and STEM 2014 show counts of industries for all three levels of High Tech intensity. For OECD (2003) we use industries classified as high and medium-High Technology and for OECD (2015a) we use high and medium-High Technology industries. Since not all major sectors have the same number of industries, Appendix G shows the employment share coverage of each method in each major sector.

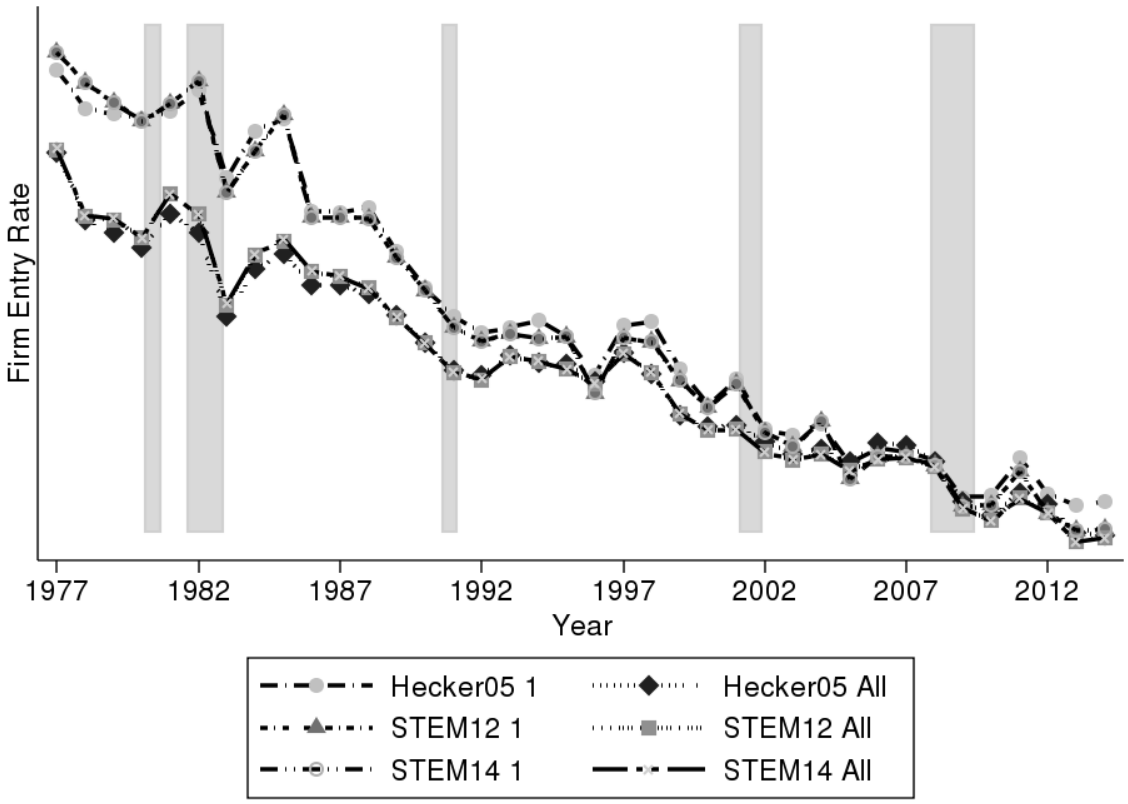
Table 7: BDS High Tech Industries

NAICS	Description
2111	Oil and Gas Extraction
3254	Pharmaceutical and Medicine Manufacturing
3341	Computer and Peripheral Equipment Manufacturing
3342	Communications Equipment Manufacturing
3344	Semiconductor and Other Electronic Component Manufacturing
3345	Navigational, Measuring, Electromedical, and Control Instruments Manufacturing
3364	Aerospace Product and Parts Manufacturing
5112	Software Publishers
5171	Wired Telecommunications Carriers
5179	Other Telecommunications
5182	Data Processing, Hosting, and Related Services
5191	Other Information Services
5413	Architectural, Engineering, and Related Services
5415	Computer Systems Design and Related Services
5417	Scientific Research and Development Services

Source: Table lists 4-digit 2007 NAICS industries in the union of industries identified by Hecker (2005) and recalculated STEM concentrations using BLS 2012 and 2014 industry-occupation employment matrices.

Figures

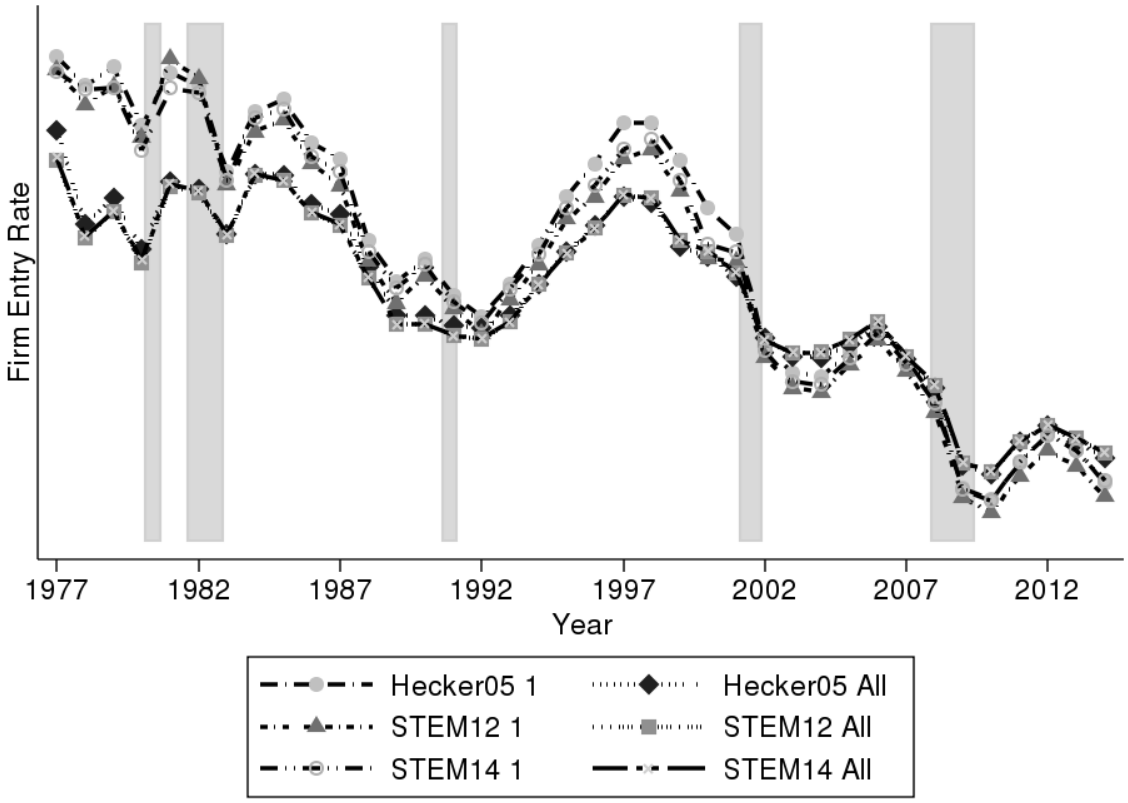
Figure 1: Manufacturing High Tech Firm Entry Rates–STEM based definitions



Source: Longitudinal Business Database, author's calculations.

Notes: Hecker05 1, STEM12 1, and STEM14 1 capture Level I industries according to Hecker 2005, STEM 2012, and STEM 2014 respectively. Hecker05 All, STEM12 All, and STEM14 All capture industries in all 3 levels according to Hecker 2005, STEM 2012, and STEM 2014 respectively. Firm entry rate calculated as 100 * the count of age zero firms at time t divided by the average number of firms in in time t and $t-1$. The y-axis has been removed to avoid the disclosure of sensitive data. Shaded regions represent recessions as defined by NBER.

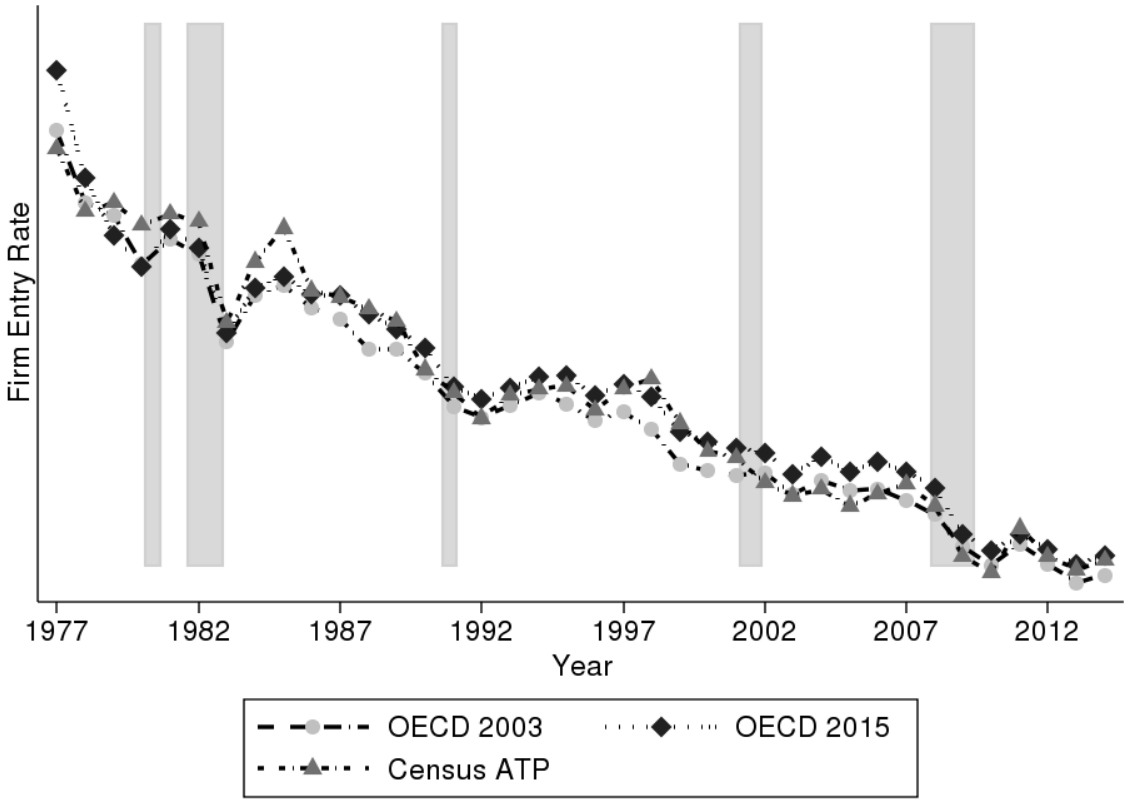
Figure 2: All High Tech Firm Entry Rates–STEM based definitions



Source: Longitudinal Business Database, author's calculations.

Notes: Hecker05 1, STEM12 1, and STEM14 1 capture Level I industries according to Hecker 2005, STEM 2012, and STEM 2014 respectively. Hecker05 All, STEM12 All, and STEM14 All capture industries in all 3 levels according to Hecker 2005, STEM 2012, and STEM 2014 respectively. Firm entry rate calculated as 100 * the count of age zero firms at time t divided by the average number of firms in in time t and $t-1$. The y-axis has been removed to avoid the disclosure of sensitive data. Shaded regions represent recessions as defined by NBER.

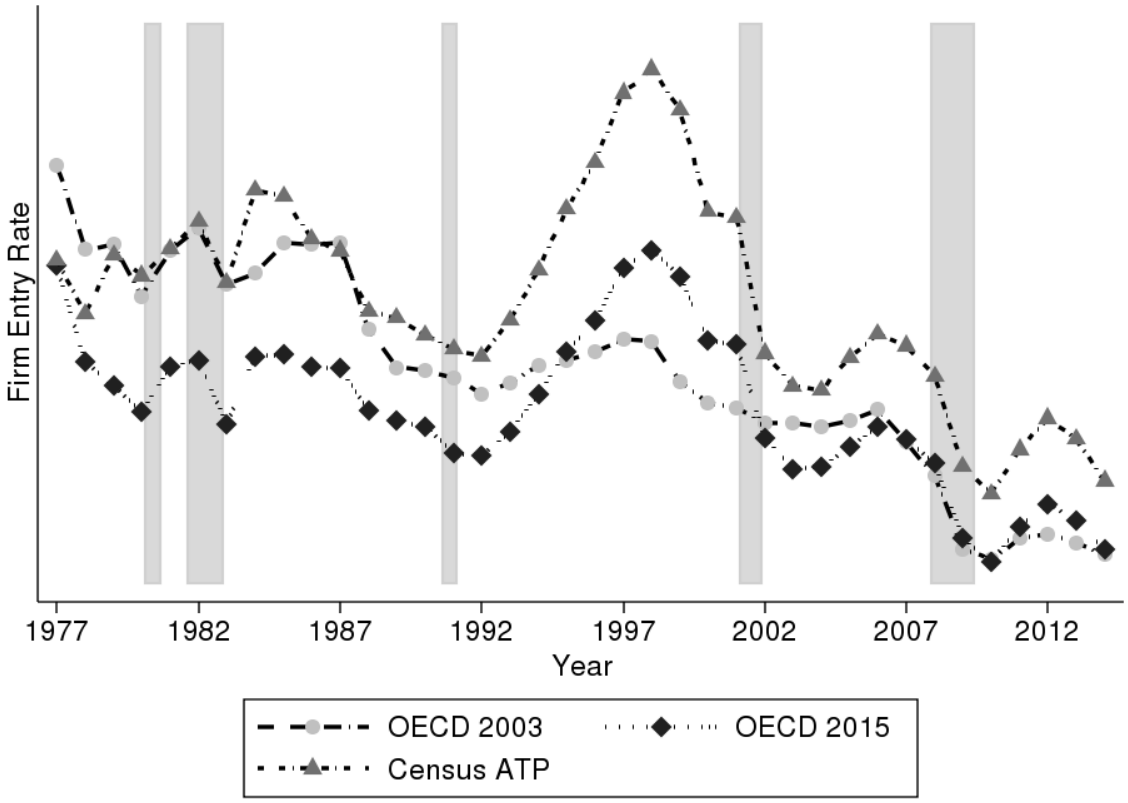
Figure 3: Manufacturing High Tech Firm Entry Rates–OECD and Census ATP definitions



Source: Longitudinal Business Database, author's calculations.

Notes: Firm entry rate calculated as 100* the count of age zero firms at time t divided by the average number of firms in in time t and $t-1$. The y-axis has been removed to avoid the disclosure of sensitive data. Shaded regions represent recessions as defined by NBER.

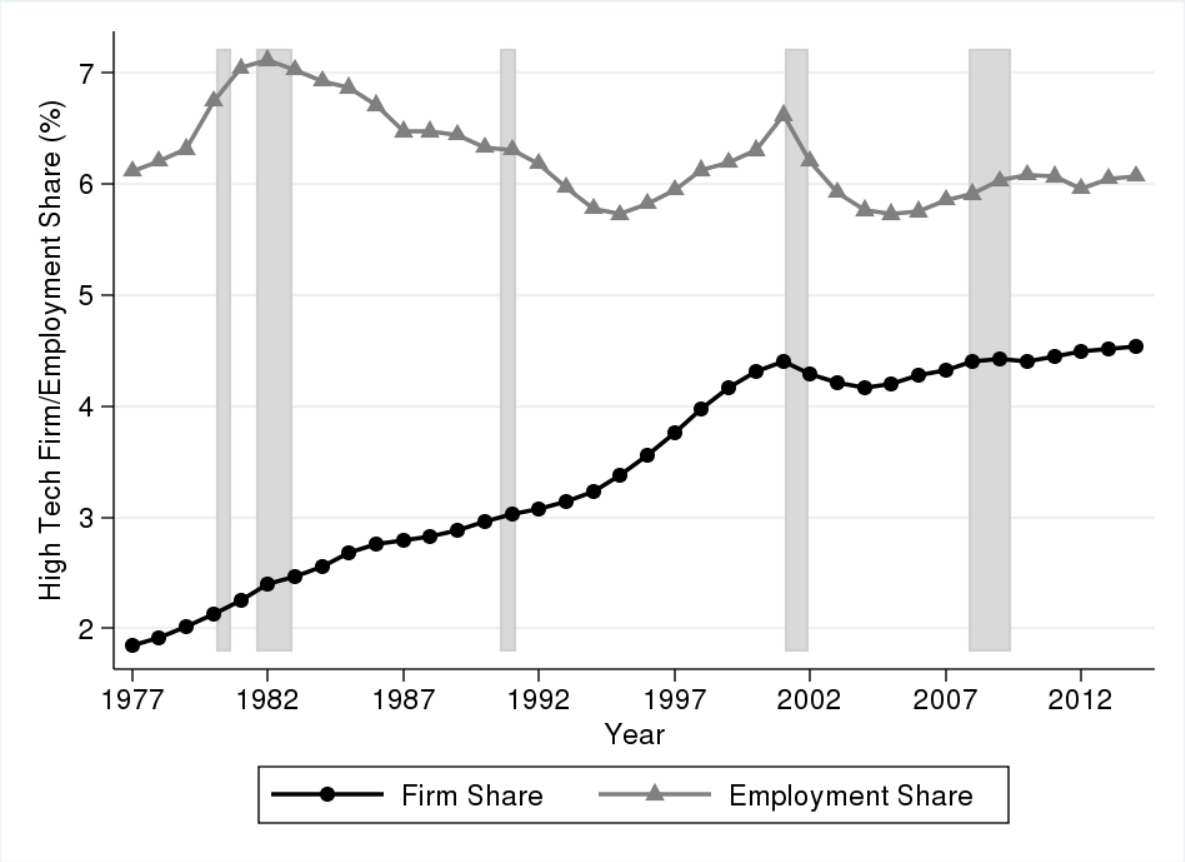
Figure 4: All High Tech Firm Entry Rates–OECD and Census ATP definitions



Source: Longitudinal Business Database, author's calculations.

Notes: Firm entry rate calculated as 100* the count of age zero firms at time t divided by the average number of firms in in time t and $t-1$. The y-axis has been removed to avoid the disclosure of sensitive data. Shaded regions represent recessions as defined by NBER.

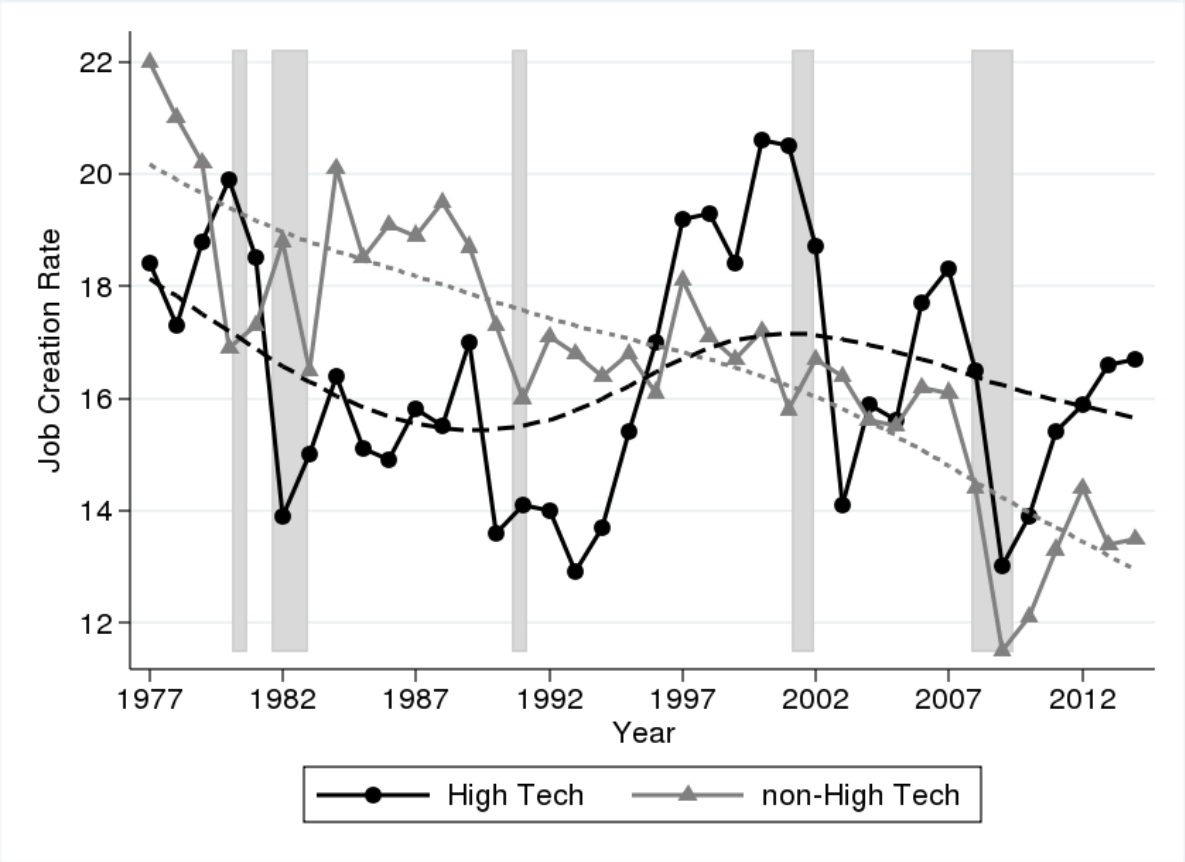
Figure 5: High Tech Firm and Employment Share



Source: BDS High Tech, author's calculations.

Notes: Y-axis does not start at zero. Firm share is calculated as the number of firms with at least one establishment in a High Tech industry divided by the total number of firms in the economy. Employment share is calculated as the sum of employment at High Tech establishments divided by the total amount of employment in the economy. Shaded regions represent recessions as defined by NBER.

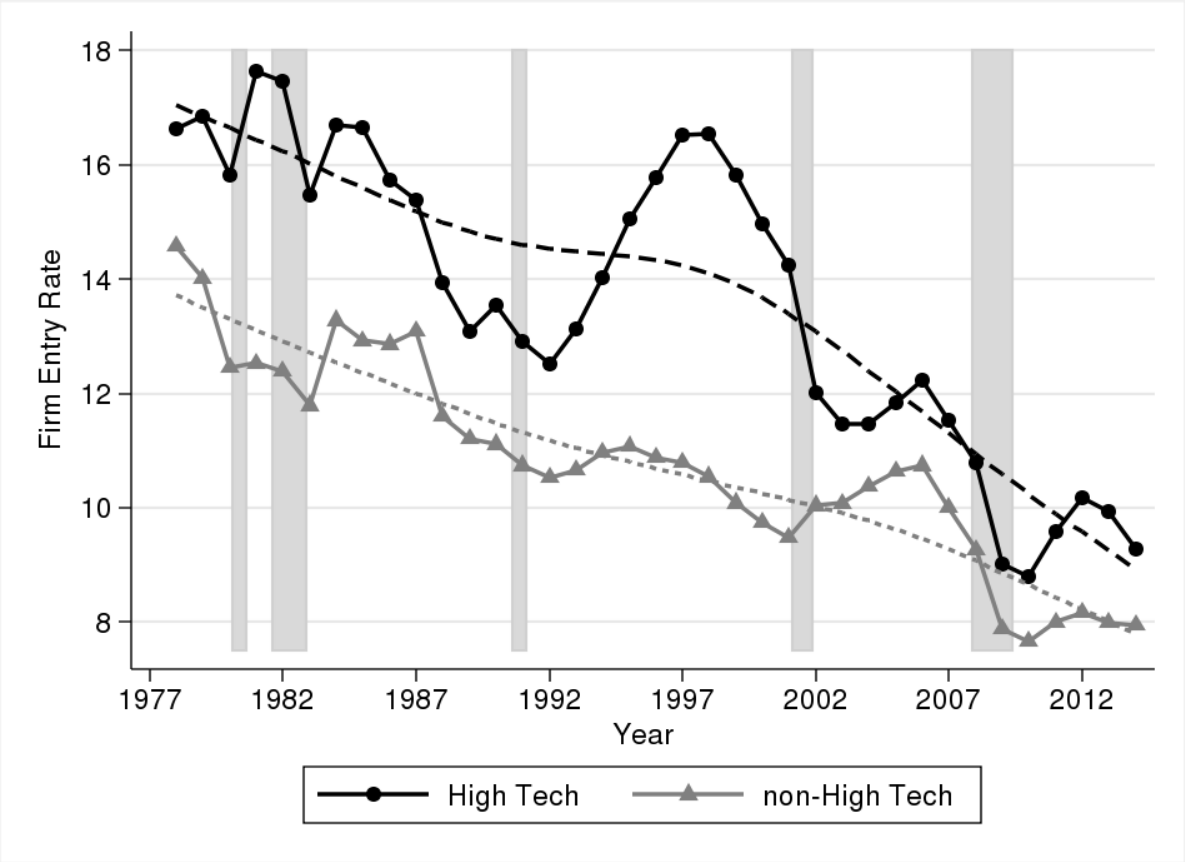
Figure 6: Job Creation Rates in High Tech and non-High Tech Industries



Source: BDS High Tech, author's calculations.

Notes: Trends calculated by applying a Hodrick-Prescott filter with a multiplier of 400. Y-axis does not start at zero. Shaded regions represent recessions as defined by NBER.

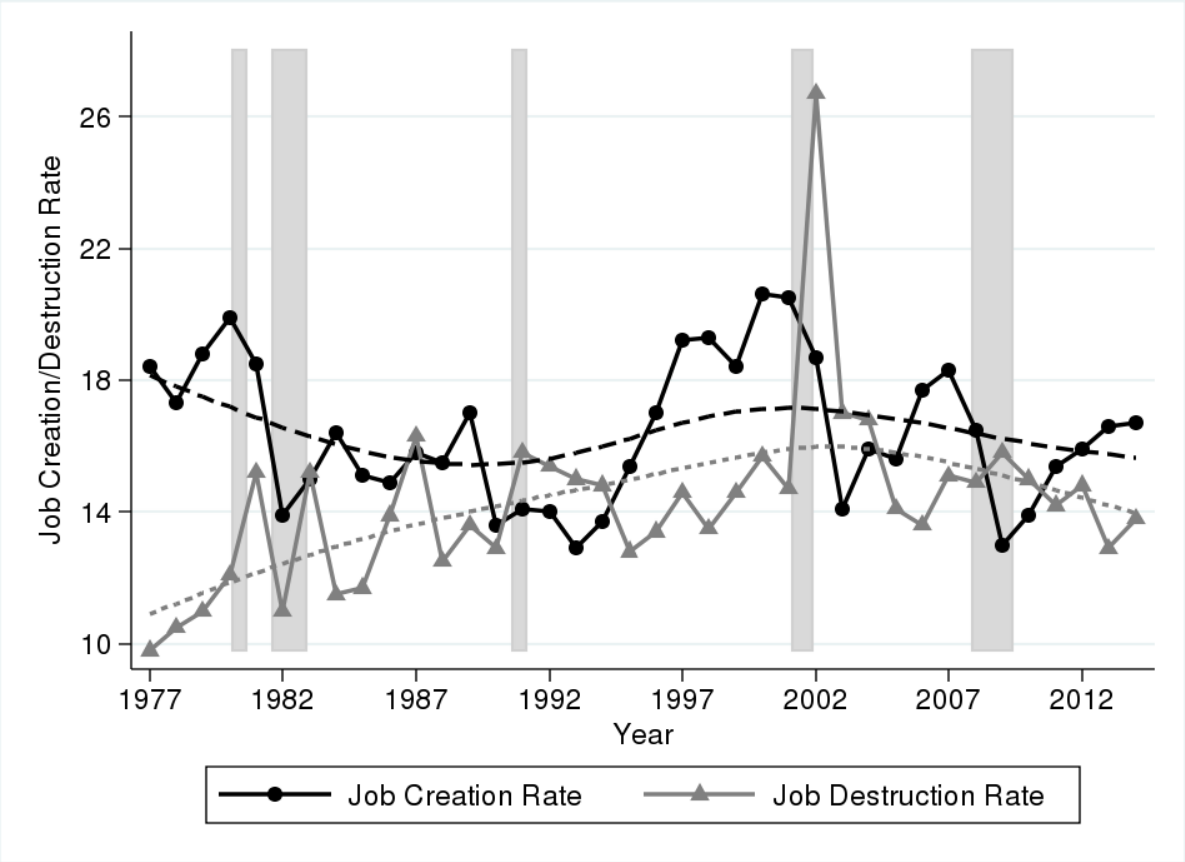
Figure 7: Firm Entry Rate High Tech vs non-High Tech



Source: BDS High Tech, author's calculations.

Notes: Trends calculated by applying a Hodrick-Prescott filter with a multiplier of 400. Firm Y-axis does not start at zero. Firm entry rate calculated as the count of age zero firms in year t divided by the average count of firms in year t and $t-1$. Shaded regions represent recessions as defined by NBER.

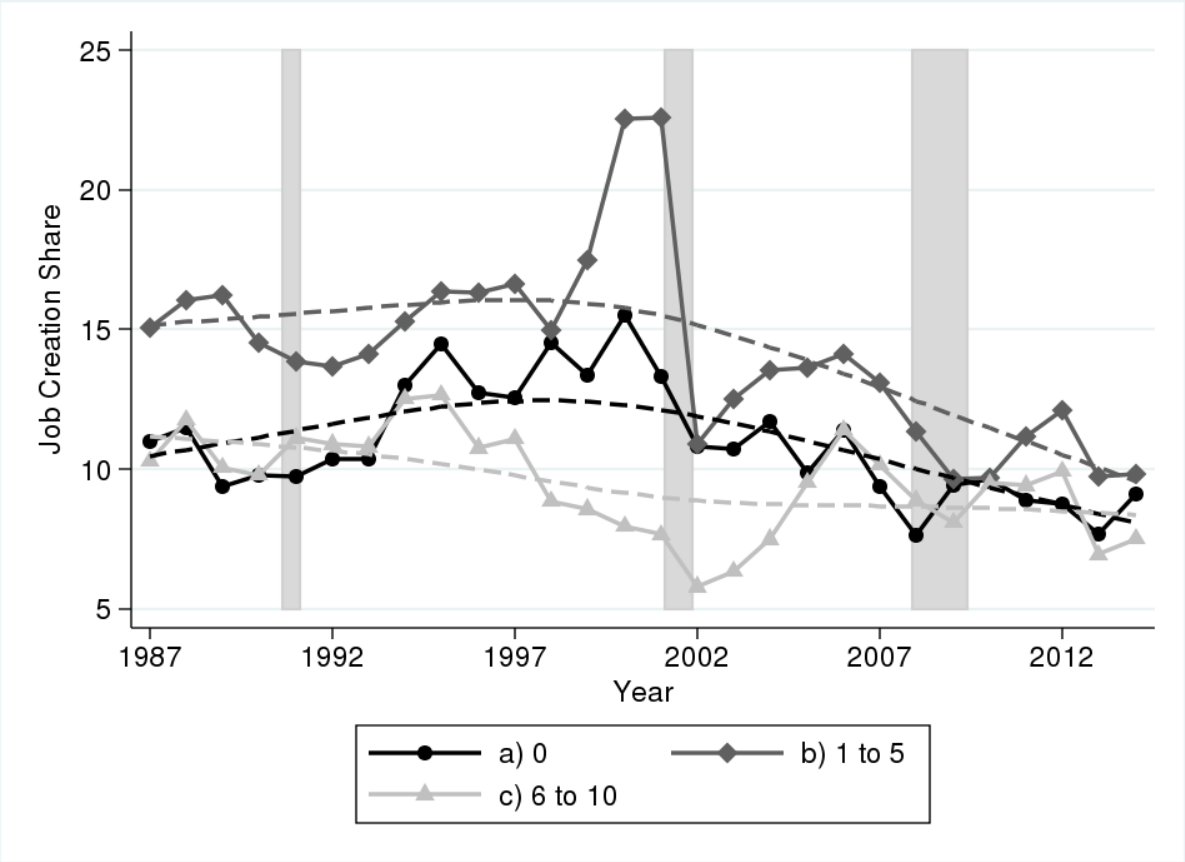
Figure 8: High Tech Job Creation Rate and Job Destruction Rate



Source: BDS High Tech, author's calculations.

Notes: Trends calculated by applying a Hodrick-Prescott filter with a multiplier of 400. Firm Y-axis does not start at zero. Shaded regions represent recessions as defined by NBER.

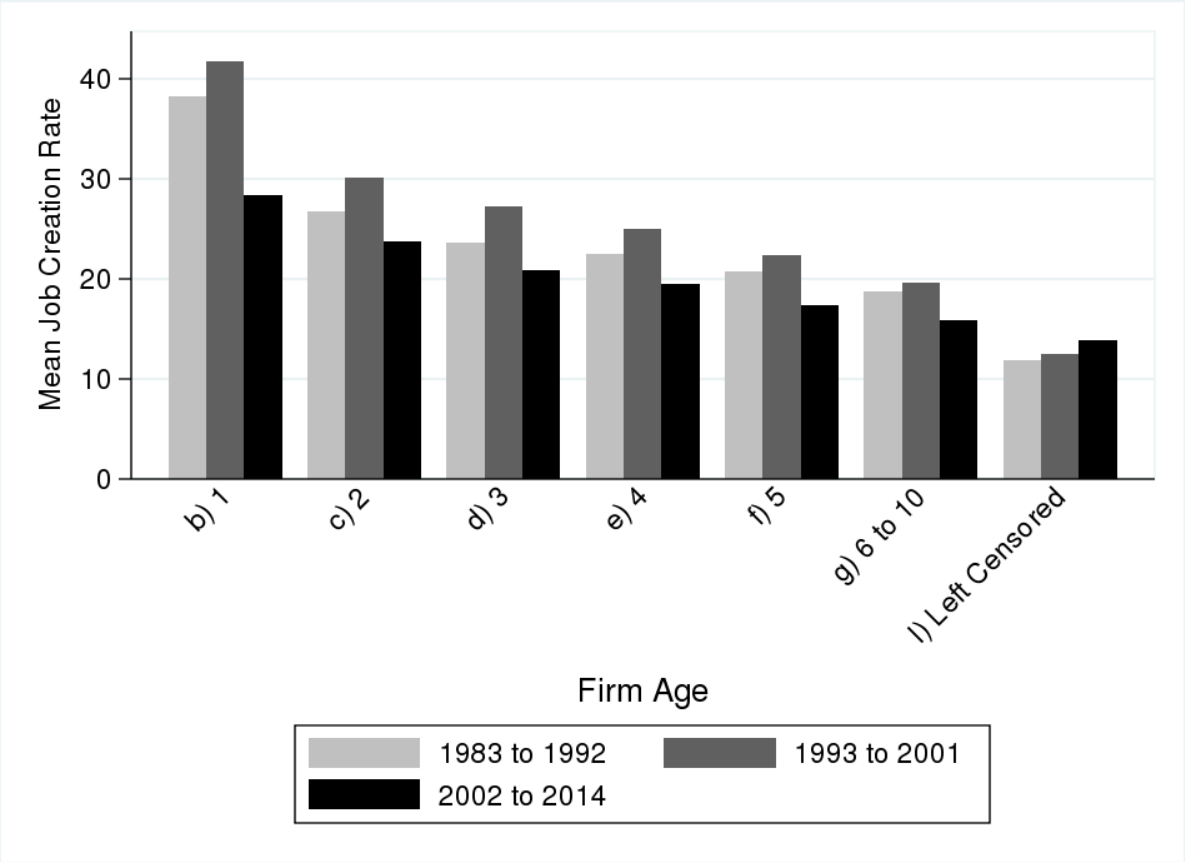
Figure 9: Job Creation Share by Firm Age



Source: BDS High Tech, author's calculations.

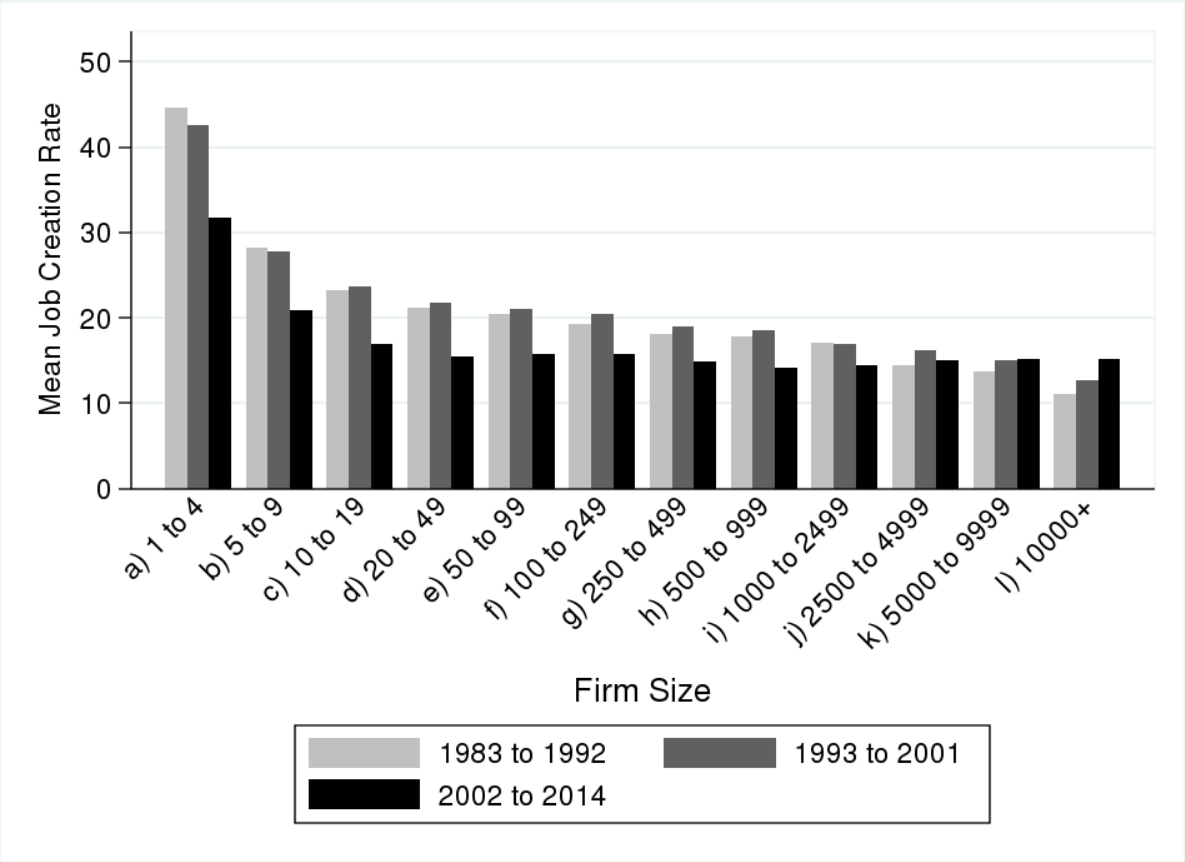
Notes: Trends calculated by applying a Hodrick-Prescott filter with a multiplier of 400. Y-axis does not start at zero. Shaded regions represent recessions as defined by NBER.

Figure 10: Average High Tech Job Creation Rates over time–By Firm Age



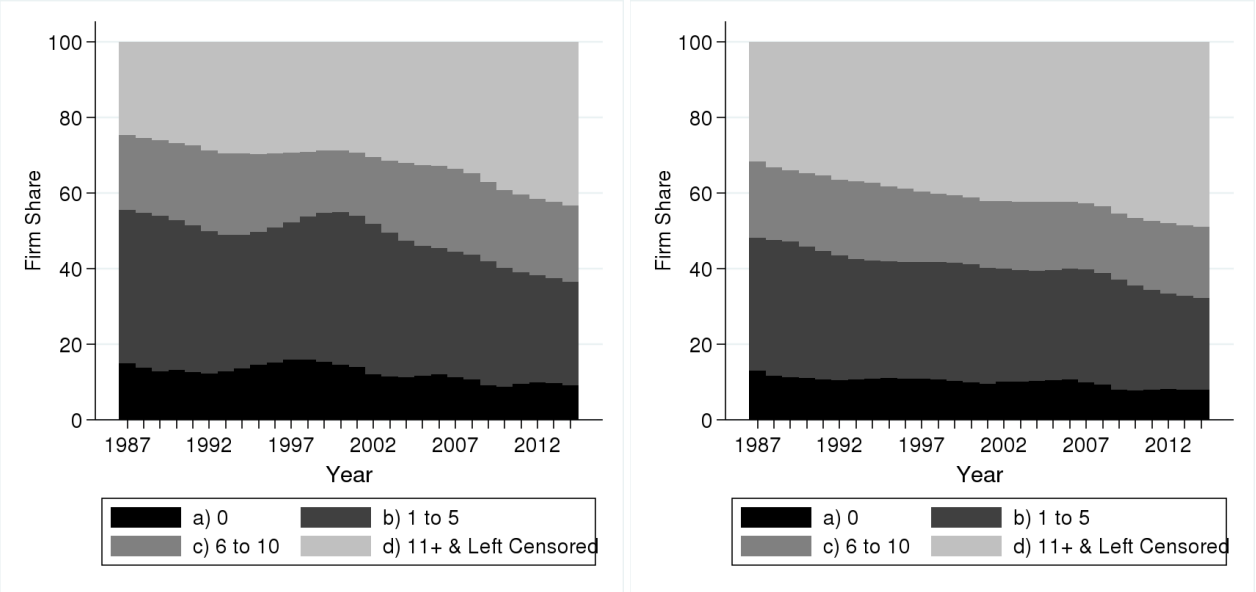
Source: BDS High Tech, author's calculations.

Figure 11: Average High Tech Job Creation Rates over time–By Firm Size



Source: BDS High Tech, author's calculations.

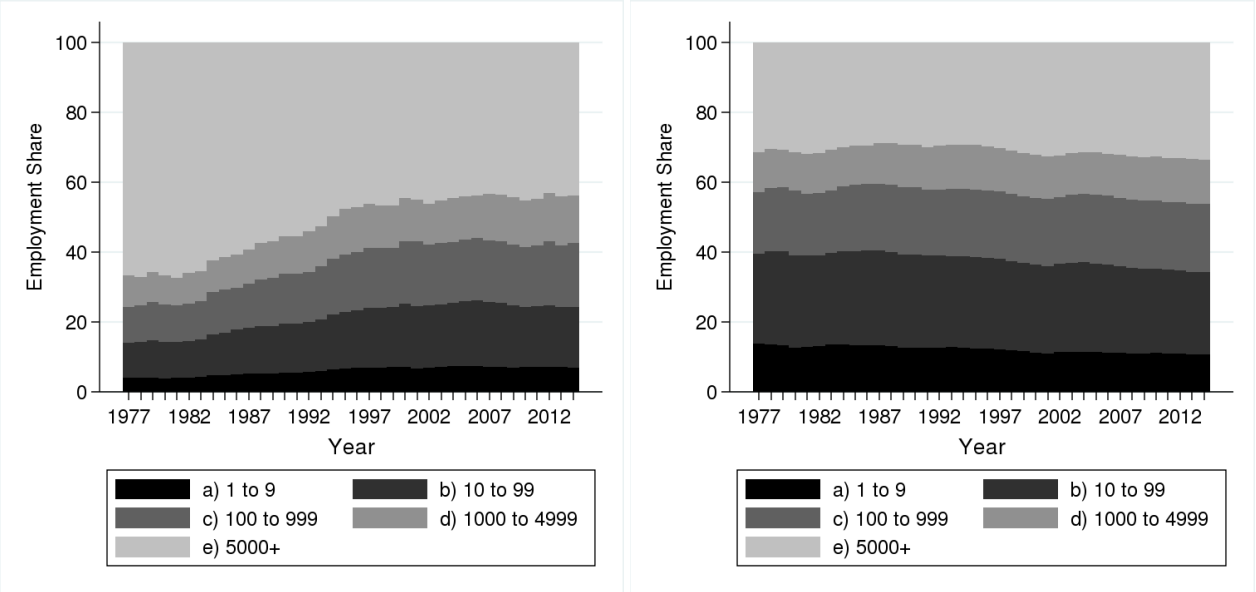
Figure 12: Firm Share by Firm Age for High Tech (left) and non-High Tech Sector (right)



Source: BDS High Tech, author's calculations.

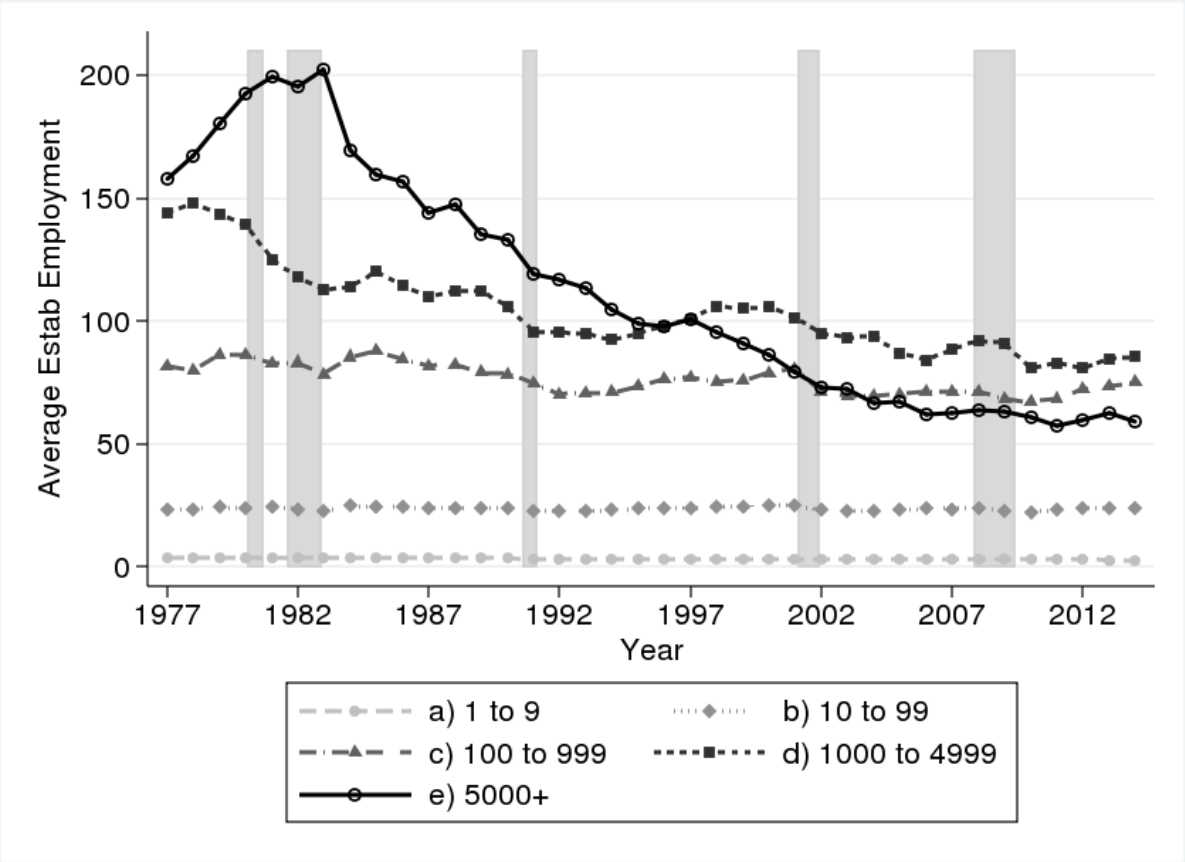
Notes: The x-axis begins at 1987 because this is the first year we observe firms in all of the listed age categories.

Figure 13: Employment Share by Firm Size for High Tech (left) and non-High Tech Sector (right)



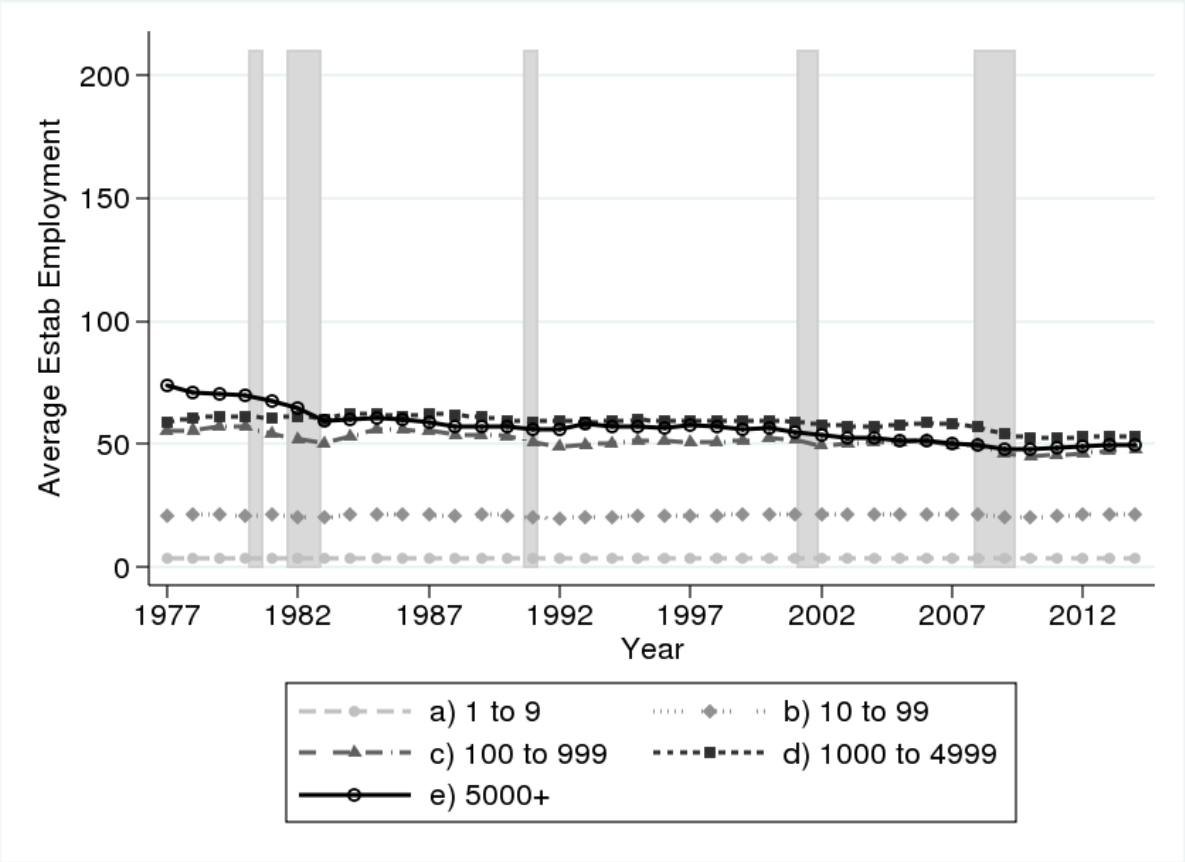
Source: BDS High Tech, author's calculations.

Figure 14: Average Establishment Size by Firm Size, High Tech



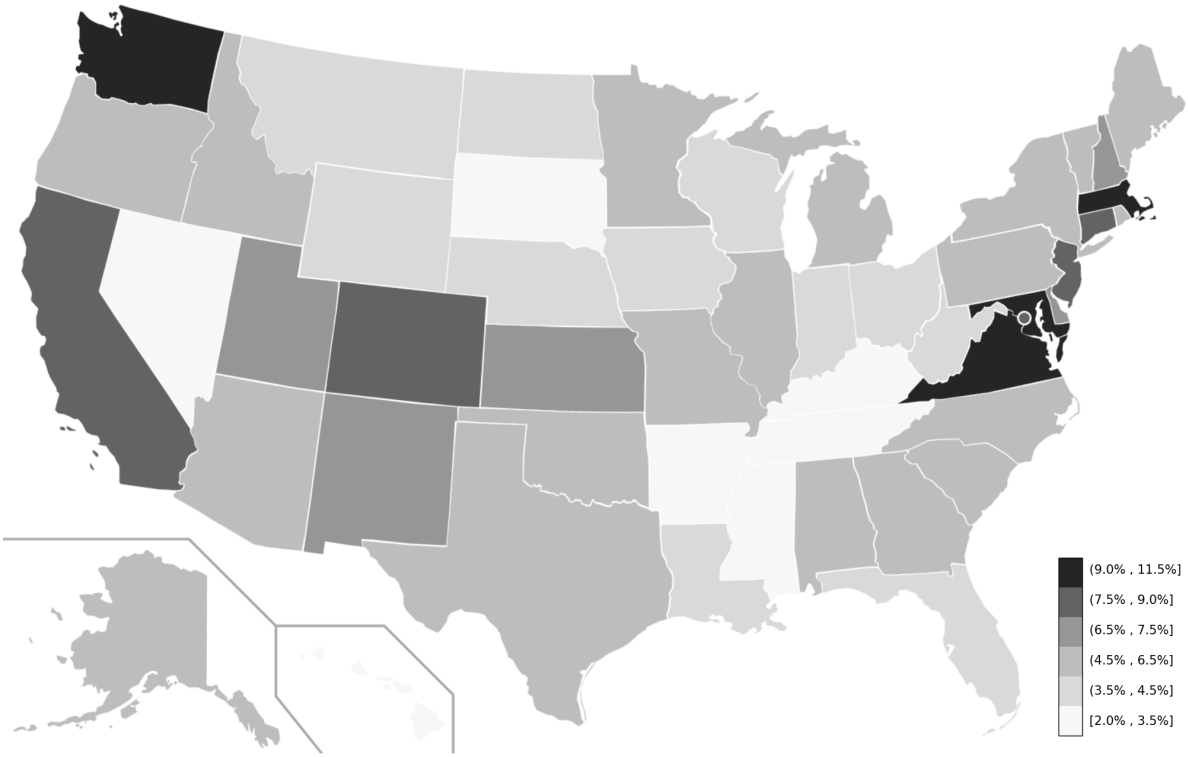
Source: BDS High Tech, author's calculations. Notes: Average establishment employment is calculated as the sum of employment in a cell divided by the total count of establishments within that cell. Shaded regions represent recessions as defined by NBER.

Figure 15: Average Establishment Size by Firm Size, non-High Tech



Source: BDS High Tech, author's calculations. Notes: Average establishment employment is calculated as the sum of employment in a cell divided by the total count of establishments within that cell. Shaded regions represent recessions as defined by NBER.

Figure 16: Average High Tech Employment Share by State, 2012-2014



Source: BDS High Tech, author's calculations.

Appendices

A Hecker (2005) High Tech Industries

NAICS	Description
Level I	
5415	Computer systems design and related services
5112	Software publishers
5413	Architectural, engineering, and related services
5417	Scientific research and development services
5181	Internet service providers and web search portals
3341	Computer and peripheral equipment manufacturing
5161	Internet publishing and broadcasting
3345	Navigational, measuring, electrometrical, and control instruments manufacturing
5182	Data processing, hosting, and related services
3364	Aerospace product and parts manufacturing
3342	Communications equipment manufacturing
3344	Semiconductor and other electronic component manufacturing
3254	Pharmaceutical and medicine manufacturing
5179	Other telecommunications
Level II	
2111	Oil and gas extraction
1131-1132	Forestry
3333	Commercial and service industry machinery manufacturing
3346	Manufacturing and reproducing magnetic and optical media
3251	Basic chemical manufacturing
4234	Professional and commercial equipment and supply merchant wholesalers
3332	Industrial machinery manufacturing
	Federal Government, excluding Postal Service
5416	Management, scientific, and technical consulting services
3343	Audio and video equipment manufacturing
2211	Electric power generation, transmission, and distribution

NAICS	Description *continued*
3252	Resin, synthetic rubber, and artificial synthetic fibers and filaments manufacturing
Level III	
5171	Wired telecommunications carriers
5511	Management of companies and enterprises
4862	Pipelines transportation of natural gass
5211	Monetary authoritiescentral bank
5172	Wireless telecommunications carriers (except satellite)
5173	Telecommunications resellers
5174	Satellite telecommunications
3353	Electrical equipment manufacturing
3259	Other chemical product and preparation manufacturing
3339	Other genera-purpose machinery manufacturing
3336	Engine, turbine, and power transmission equipment manufacturing
3255	Paint, coating, and adhesive manufacturing
3241	Petroleum and coal products manufacturing
8112	Electronic and precision equipment repair and maintenance
5612	Facilities support services
3253	Pesticides, fertilizer, and other agricultural chemical manufacturing
5232	Securities and commodity exchanges
4861	Pipeline transportation crude oil
3369	Other transportation equipment manufacturing
4869	Other pipeline transportation

Source: Hecker (2005) Table 4, pp. 64. 2002 NAICS codes.

B Hecker (1999) High Tech Industries

NAICS	Description
32411	Petroleum refineries
3251	Basic chemical manufacturing
3252	Resin, synthetic rubber, and artificial and synthetic fibers and filaments manufacturing
3253	Pesticide, fertilizer, and other agricultural chemical manufacturing
3254	Pharmaceutical and medicine manufacturing
3255	Paint, coating, and adhesive manufacturing
3256	Soap, cleaning compound, and toilet preparation manufacturing
3259	Other chemical product and preparation manufacturing
332992	Ordnance & accessories manufacturingsmall arms ammunition manufacturing
332993	Ordnance & accessories manufacturingammunition (except small arms) manufacturing
332994	Ordnance & accessories manufacturingsmall arms manufacturing
332995	Ordnance & accessories manufacturingother ordnance and accessories manufacturing
3331	Agriculture, construction, and mining machinery manufacturing
3332	Industrial machinery manufacturing
3333	Commercial and service industry machinery manufacturing
3336	Engine, turbine, and power transmission equipment manufacturing
3339	Other general purpose machinery manufacturing
3341	Computer and peripheral equipment manufacturing
3342	Communications equipment manufacturing
3343	Audio and video equipment manufacturing
3344	Semiconductor and other electronic component manufacturing
3345	Navigational, measuring, electromedical, and control instruments manufacturing
3346	Manufacturing and reproducing magnetic and optical media
3353	Electrical equipment manufacturing
33599	All other electrical equipment and component manufacturing
3361	Motor vehicle manufacturing
3362	Motor vehicle body and trailer manufacturing

NAICS	Description *continued*
3363	Motor vehicle parts manufacturing
3364	Aerospace product and parts manufacturing
3391	Medical equipment and supplies manufacturing
5112	Software publishers
514191	On-line information services
5142	Data processing services
5413	Architectural, engineering, and related services
5415	Computer systems design and related services
5416	Management, scientific, and technical consulting services
5417	Scientific research and development services
6117	Educational support services
811212	Computer and office machine repair and maintenance

Source: NSF (2006).
Notes: Science and Engineering Indicators. Chapter 8 State Indicators, Technical Note: "Defining High-Technology Industries." (http://www.nsf.gov/statistics/seind06/c8/c8.cfm?opt=9) accessed 11/19/2015, original list of SIC industries Hecker (1999) Table 1, pp 20. This table captures a joint effort by the Office of Technology Policy and the Census Bureau to convert the SIC codes found in Hecker (1999) to the 1997 edition of NAICS.

C High Tech Industries by R&D Intensity

ISIC Rev. 3 Code	Description
High Tech Industries	
353	Aircraft and spacecraft
2423	Pharmaceuticals
30	Office, accounting, and computing machinery
32	Radio, TV, and communications equipment
33	Medical, precision, and optical instruments
Medium Tech Industries	
31	Electrical machinery and apparatus, n.e.c.
34	Motor vehicles, trailers and semi-trailers
24 excl. 2423	Chemicals excluding pharmaceuticals
352 + 359	Railroad equipment and transport equipment, n.e.c.
29	Machinery and equipment, n.e.c.
Medium-Low Tech Industries	
351	Building and repairing of ships and boats
25	Rubber and plastics products
23	Coke, refined petroleum products and nuclear fuel
26	Other non-metallic mineral products
27-28	Basic metals and fabricated metal products
Low Tech Industries	
36-37	Manufacturing, n.e.c.; Recycling
20-22	Wood, pulp, paper, paper products, printing and publishing
15-16	Food products, beverages, and tobacco
17-19	Textiles, textile products, leather, and footwear

Source: OECD (2003), Annex 1.1.

D Advanced Tech Products High Tech Industries

Major Tech	NAICS	Description
Biotechnology	3254	Pharmaceutical and medicine manufacturing
	5417	Scientific R&D services
Life science technologies	3345	Navigational, measuring, electromedical, and control instrument manufacturing
	3254	Pharmaceutical and medicine manufacturing
Optoelectronics	3341	Computer and peripheral equipment manufacturing
	3344	Semiconductor and other electronic component manufacturing
Information and communications	3341	Computer and peripheral equipment manufacturing
	3342	Communications equipment manufacturing
	3345	Navigational, measuring, electromedical, and control instrument manufacturing
	5112	Software publishing
	5415	Computer systems design and related services
Electronics	3344	Semiconductor and other electronic component manufacturing
Flexible manufacturing	3332	Industrial machinery manufacturing
	3335	Metalworking machinery manufacturing
Advanced materials	3344	Semiconductor and other electronic component manufacturing
	3346	Manufacturing and reproducing magnetic and optical medial
	3359	Other electrical equipment and component manufacturing
Aerospace	3364	Aerospace product and parts manufacturing
Weapons	3329	Other fabricated metal products manufacturing
Nuclear technology	3324	Boiler, tank, and shipping container manufacturing
	3251	Basic chemical manufacturing

Source: Hecker (2005) Exhibit 1, pp 67. Notes: 2002 NAICS codes.

E High Tech Products High Tech Industries

Sector	SITC Rev. 3 Code
Aerospace	[7921+7922+7923+7924+7925+79293+(714-71489-71499)+87411]
Computers-office machines	[75113+75131+75132+75134+(752-7529)+75997]
Electronics-telecommunications	[76381+76383+(764-76493-76499) +7722+77261+77318 +77625+7763+7764+7768+89879]
Pharmacy	[5413+5415+5416+5421+5422]
Scientific instruments	[774+8711+8713+8714+8719+87211+(874-87411-8742)+88111+88121+88411+88419+89961+89963++89967]
Electrical machinery	[77862+77863+77864+77865+7787+77844]
Chemistry	[52222+52223+52229+52269+525+57433+591]
Non-electrical machinery	[71489+71499+71871+71877+72847+7311+73131+73135 585716 +73733+73735]
Armament	[891–]

Source: Hatzichronoglou (1997) Table 2.

F Industries Leaving/Entering STEM Classification

NAICS	Description
2012 Only	
336510	Railroad Rolling Stock Manufacturing
2014 Only	
212210	Iron Ore Mining
212221	Gold Ore Mining
212222	Silver Ore Mining
212231	Lead Ore and Zinc Ore Mining
212234	Copper Ore and Nickel Ore Mining
212291	Uranium-Radium-Vanadium Ore Mining
212299	All Other Metal Ore Mining

Source: List of industries that only appear as High Tech in the 2012 or 2014 vintage of the OES STEM concentration methodology. Note that STEM concentrations are calculated at the 4-digit NAICS level. One 4-digit industry was found only in the 2012 STEM concentration data, 3365, which in turn has only one associated 6-digit industry. One 4-digit industry was found only in the 2014 STEM concentration data, 2122, which in turn has 7 associated 6-digit industries.

G Employment Share Coverage by Major Sector

Major Sector	Hecker (2005)	STEM 2012	STEM 2014	Census ATP	OECD (2003)	OECD (2015)
Agriculture, Forestry, Fishing and Hunting	2.8	0	0	0	0	0
Arts & Accommodation	0	0	0	0	0.1	0
Construction	0	0	0	0	0	0
Education & Health	0	0	0	0	6.2	0
FIRE	0.2	16.6	16.6	0	82	0
Information	50	52	52	12.8	61.4	13.2
Management & Support	25.7	24.3	24.3	0	87.4	0
Manufacturing	22.7	22.2	22	19.7	26.2	47.3
Mining & Utilities	48.9	48.9	51.5	0	17.8	0
Retail & Wholesale	3.5	3.5	3.5	0	0	0
Services	30.1	29.1	29.1	10.4	60.4	10.7
Transport & Warehousing	1	0.8	0.2	0	16.6	0

Notes: Percent of employment within major sectors covered by High Tech industries according to each methodology. Employment counts are derived from 2007 count business pattern 6-digit industry level aggregates (http://www.census.gov/data/datasets/2007/econ/cbp/2007-cbp.html accessed 10/17/2016). For Hecker (2005), STEM 2012, and STEM 2014 show counts of industries for all three levels of High Tech intensity. For OECD (2003) we use industries classified as high and medium-High Technology and for OECD (2015) we use high and medium-High Technology industries.